The pleiotropic transcriptional regulator NlpR contributes to the modulation of nitrogen metabolism, lipogenesis and triacylglycerol accumulation in oleaginous rhodococci

Martín A. Hernández[1], Julia Lara[2], Gabriela Gago[2], Hugo Gramajo[2], Héctor M. Alvarez[1*]

[1]*INBIOP (Instituto de Biociencias de la Patagonia), Consejo Nacional de Investigaciones Científicas y Técnicas, Facultad de Ciencias Naturales, Universidad Nacional de la Patagonia San Juan Bosco, Ruta Provincial N° 1, Km 4-Ciudad Universitaria 9000, Comodoro Rivadavia, Chubut, Argentina.*

[2]*IBR (Instituto de Biología Molecular y Celular de Rosario), Consejo Nacional de Investigaciones Científicas y Técnicas. Facultad de Ciencias Bioquímicas y Farmacéuticas, Universidad Nacional de Rosario, Ocampo y Esmeralda 2000, Rosario, Santa Fe, Argentina.*

Running title: *Regulation of triacylglycerol accumulation in rhodococci*

Key words: *Rhodococcus*, triacylglycerols, lipogenesis, regulation

*Corresponding author: Héctor M. Alvarez, Instituto de Biociencias de la Patagonia, Consejo Nacional de Investigaciones Científicas y Técnicas, Facultad de Ciencias Naturales (CONICET), Universidad Nacional de la Patagonia San Juan Bosco, Ruta Provincial N° 1, Km 4-Ciudad Universitaria 9000, Comodoro Rivadavia, Chubut, Argentina. Tel.: +54 297 4210212; E-mail: halvarez@unpata.edu.ar

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as an 'Accepted Article', doi: 10.1002/mmi.13564

Summary

The regulatory mechanisms involved in lipogenesis and triacylglycerol (TAG) accumulation are largely unknown in oleaginous rhodococci. In this study we identified a regulatory protein (here called NlpR: Nitrogen lipid Regulator), which contributes to the modulation of nitrogen metabolism, lipogenesis and triacylglycerol accumulation in oleaginous rhodococci. Under nitrogen deprivation conditions, in which TAG accumulation is stimulated, the *nlpR* gene was significantly upregulated, whereas a significant decrease of its expression and TAG accumulation occurred when cerulenin was added. The *nlpR* disruption negatively affected the nitrate/nitrite reduction as well as lipid biosynthesis under nitrogen-limiting conditions. In contrast, its overexpression increased TAG production during cultivation of cells in nitrogen-rich media. A putative "NlpR-binding motif" upstream of several genes related to nitrogen and lipid metabolisms was found. The *nlpR* disruption in RHA1 strain led to a reduced transcription of genes involved in nitrate/nitrite assimilation, as well as in fatty acid and TAG biosynthesis. Purified NlpR was able to bind to *narK*, *nirD*, *fasI*, *plsC* and *atf*3 promoter regions. We suggest that NlpR acts as a pleiotropic transcriptional regulator by activating of nitrate/nitrite assimilation genes and others genes involved in fatty acid and TAG biosynthesis, in response to nitrogen deprivation.

Introduction

In the recent years, oleaginous microorganisms have attracted great interest because of their biotechnological potential as single cell oils (SCOs) sources with application in the industry. SCOs have emerged as a promising source of oleochemicals, food or cosmetic additives and biofuels (Alvarez, 2010). Oleaginous actinobacteria belonging to *Rhodococcus* genus are promising candidates for industrial purposes. Some members of this genus, such as *Rhodococcus opacus* PD630 and *Rhodococcus jostii* RHA1, are able to produce up to 60% (cellular dry weight, CDW) of triacylglycerols (TAG) under nitrogen-limiting conditions from a wide range of carbon sources including sugars, hydrocarbons, or organic wastes (Alvarez *et al.*, 1996; Hernández *et al.*, 2008). The TAG biosynthesis in rhodococci is a complex process that involves several catalytic pathways, key cofactors/precursors and probably a specific regulatory circuit to coordinate those metabolic processes occurring in the cell (Alvarez, 2016). Holder *et al.*, (2011) reported at least 261 genes involved in the *R. opacus* PD630 TAG cycle based on metabolic reconstruction and gene family analysis. On the other hand, Chen *et al.*, (2013) described at least 177 genes associated with lipid metabolism in the same strain based on transcriptome and lipid-droplet proteome analyses. Further, in a previous integrative proteome study, we reported the significant upregulation of several proteins during TAG accumulation when cells of *R. jostii* RHA1 were grown under nitrogen-limiting conditions (Dávila Costa *et al.*, 2015). In the last years, several genes/proteins and their contributions on TAG metabolism have been well described in *R. opacus* PD630 or *R. jostii* RHA1 used as models. Examples of these studies include proteins such as diacylglycerol acyltransferase enzymes (DGAT) (Alvarez et al., 2008; Hernández et al., 2013, Amara et al. 2016), a NADPH-dependent glyceraldehyde-3-phosphate dehydrogenase (TadD) (MacEachran and Sinskey, 2013), a phosphatidic acid phosphatase type 2 enzyme (PAP2) (Hernández et al., 2015), an ATP-binding cassette transporter (Lpt1) (Villalba and Alvarez, 2014) and structural proteins involved in lipid body ontogeny (TadA) (MacEachran et al., 2010; Ding et al., 2012). Besides enzymes, transporters and structural proteins, the TAG-accumulating machinery of rhodococci may include an unknown regulatory network probably integrated by global and specific regulatory proteins, which could coordinate the transition to a TAG-producing phenotype. The sole regulatory protein involved in actinobacterial lipid metabolism reported so far is FasR, which controls gene expression of fatty acid biosynthesis (FabDHPF operon) in *Streptomyces coelicolor* (Arabolaza *et al.*, 2010). In

contrast, the mechanisms controlling genes for lipogenesis in oleaginous rhodococci are largely unknown. The knowledge of this regulatory circuit in oleaginous rhodococci may be also useful for improving oil yields for biotechnological purposes. For these reasons, we decided to explore the role of some transcriptional regulatory proteins that are significantly upregulated during TAG accumulation in oleaginous rhodococci. After analyzing different "omics" studies in *R. jostii* RHA1 and *R. opacus* PD630, we observed a reproducible induction of a putative transcriptional regulator coded by the *RHA1_RS31140* (formerly RHA1_*ro06368*) gene under nitrogen-limiting conditions (Chen *et al.*, 2013; Dávila Costa *et al.*, 2015; Amara *et al.*, 2016). This protein was always negligible under nitrogen-rich conditions, which promote cell growth but not TAG accumulation.

According bioinformatic data, RHA1_RS31140 is ortholog of NnaR (nitrate nitrite Regulator), a transcriptional regulator involved in the uptake and reduction of NO_3^{-}/NO_2^{-} in *Streptomyces coelicolor*. Under the control of GlnR, a global regulator of nitrogen metabolism, NnaR acts as a co-activator of GlnR to control the genes involved in uptake and reduction of NO_3^{-}/NO^{2}- in that bacterium (Amin *et al.*, 2012). GlnR mutants of *Mycobacterium smegmatis* showed a significant decreased expression of *msmeg_0432* gene (*nnaR* ortholog) in response to nitrogen starvation, confirming its occurrence as part of the GlnR regulon also in mycobacteria (Jeßberger *et al.*, 2013; Jenkins *et al.*, 2013).

Since nitrogen limitation is a requirement for triggering lipogenesis and TAG accumulation in oleaginous rhodococci, a putative regulatory mechanism linking nitrogen and lipid metabolisms was expected. In this study, we found that the transcriptional regulator encoded by the *RHA1_RS31140* gene (here called NlpR, for Nitrogen-lipid Regulator), plays an important role in the crosstalk between nitrogen metabolism and lipid metabolic pathways in *R. jostii* RHA1. This is the first regulatory protein identified and characterized, contributing to the TAG accumulation in oleaginous rhodococci. This finding provides new insights into the regulatory network and its components which control lipogenesis in these microorganisms.

Results

The putative NlpR regulator is only present in actinomycetes

The putative transcriptional regulator NlpR is coded by the *RHA1_RS31140* (formerly *ro06368*) and *opag_03371* (PD630_LPD03034) genes in *R. jostii* RHA1 and *R. opacus* PD630, respectively. The predicted NlpR proteins possess two conserved domains: The N-terminal region containing an enzyme-like domain similar to uroporphyrinogen-III synthases (HemD), which catalyzes the uroporphyrinogen III synthesis from hydroxymethylbilane and as part of the pyrrole-containing compounds biosynthesis (Stamford *et al.,* 1995), and the C-terminal region similar to the DNA binding domains of OmpR-like response regulators. Full protein sequence alignment by BLAST algorithm with all bacterial proteins available in database, showed the occurrence of NlpR homologs in all analyzed *Rhodococcus* species as well as in other actinobacteria belonging to *Nocardia*, *Gordonia*, *Mycobacterium*, *Streptomyces* and *Amycolatopsis* genera, with significant coverage percentages (>95%) and identities (I >50%) (Fig. 1 and Fig. S1). NlpR sequences from oleaginous rhodococcal species, such as *R. opacus*, *R. jostii* RHA1, and *R. wratislaviensis*, were almost identical and formed a tight sub-cluster (Fig. S2). In contrast, full NlpR protein was not detected in other bacteria beyond actinobacteria. Similar results were obtained by using the last 87 amino acids from the C-terminal DNA binding domain as query in the alignments analyses (aa 299-385 in NlpR of RHA1 strain). Again, NlpR homologs were found only among bacteria belonging to the actinomycetes group (data not shown). Interestingly, full NlpR protein was not detected in bacteria belonging to the *Corynebacterium* genus.

On the other hand, the structural prediction of NlpR based on bioinformatic analyses revealed similarities with known structures for the N-terminal and C-terminal regions, respectively. The N-terminal region of NlpR assembled (confidence of 100% and identity of 30%) with an uroporphyrinogen III synthase of *Thermus thermophiles*. The C-terminal region assembled with domains of PhoB-like proteins associated to DNA binding (confidence of 99.8-99.9% and identity of 24-26%).

The *nlpR* gene is located within a cluster containing the *narK*, *nirB* and *nirD* genes, associated with nitrate/nitrite uptake and assimilation processes in *R. jostii* RHA1 (Fig. 2). Interestingly, this cluster arrangement is highly conserved in all analyzed rhodococcal species, whereas in the other related

actinobacteria, the *nlpR* homologs genes are also often close to the *narK*, *nirB* or *nirD* genes (Fig. 2). This arrangement suggests a similar functional role of *nlpR* gene in *R. jostii* and *R. opacus* related to nitrogen metabolism as described in *Streptomyces coelicolor* (Amin *et al.*, 2012), and particularly in the modulation of uptake and assimilation processes of NO_3^- and NO_2^-.

NlpR coding gene is highly expressed under nitrogen-limiting conditions

Recently, we performed an integrated proteome study in the oleaginous *R. jostii* RHA1 to compare the protein abundance during cultivation in media with and without nitrogen source and with gluconate as sole carbon source. This study revealed that the complete cluster *narK-nirB-nirD-nlpR* was significantly upregulated during cultivation of cells under nitrogen-limiting conditions in comparison with those cells cultivated in nitrogen-rich medium (Dávila Costa *et al.*, 2015). Similar results were observed after analysis of the transcriptomic data reported by Chen *et al.*, (2013) for the oleaginous *R. opacus* strain PD630, where the ortholog *opag_03371* gene (also annotated as Pd630_LPD03034) and the adjacent genes *nirB* and *nirD* were highly upregulated under nitrogen-limiting conditions as compared to nitrogen-rich conditions.

In order to confirm these results, we performed here a semi-quantitative expression analysis of *nlpR* gene by RT-PCR with RHA1 and PD630 cells cultivated in both nitrogen-free and -rich conditions. Our results confirmed the upregulation of *nlpR* gene during cultivation of cells under nitrogen deprivation conditions in both RHA1 and PD630 strains (Fig. 3B). As is shown in Fig. 3A, the induction of the *nlpR* genes in these strains correlated with a high TAG accumulation under those conditions.

NlpR coding gene is repressed under nitrogen-limiting conditions in the presence of cerulenin

Previously, we observed a good correlation between the upregulation of NlpR and the TAG biosynthesis and accumulation in *R. jostii* and *R. opacus*. In order to establish a functional relationship between these events, we analyzed the effect of cerulenin (Fig. 3A), a potent inhibitor of *de novo* fatty acid biosynthesis and concomitantly of TAG accumulation (Alvarez 1997; Hernández *et al.*, 2010), on the expression of the studied transcriptional regulator. The effect of cerulenin on the expression of *nlpR* gene was monitored by semiquantitative RT-PCR analysis in *R. jostii* and *R. opacus* cultivated under nitrogen deprivation conditions. Interestingly, the addition of

cerulenin to the nitrogen-deprived media promoted a decrease in the *nlpR* transcription level in both RHA1 and PD630 strains (Fig. 3C). These results suggested a functional relationship between NlpR and the TAG metabolism in these oleaginous rhodococci.

Disruption of nlpR gene affects cell growth with NO_3^-, NO_2^- and low concentrations of $NH4^+$

Previous studies demonstrated that neighbor genes upstream of *nlpR* participate in nitrate/nitrite assimilation as alternative nitrogen sources in *R. jostii* RHA1 (Iino *et al.*, 2012; Iino *et al.*, 2013). To analyze the functional relationship of NlpR with nitrate/nitrite reduction, *nlpR* gene of strain RHA1 was disrupted and the phenotype of the resulting mutant strain (RHA1::*nlpR*) was analyzed under nitrogen-rich and -poor conditions. There was no difference in growth profile between RHA1::*nlpR* and the wild type (WT) strain (RHA1) during cultivation of cells in nitrogen-rich media (LB and MSM1, which contains 1 g/L of NH_4Cl) (Fig. 4A). NlpR, which was only poorly expressed in nitrogen-rich media (Fig. 3B), seemed not to play a relevant role under these conditions. Similar results were observed when urea was used as nitrogen source at 1 g/L in culture media (data not shown).

On the other hand, a slight decrease of cellular growth was observed when cells of RHA1::*nlpR* were cultivated with ammonium chloride as sole nitrogen source at low concentrations (MSM0.1) (Fig. 4B). When NO_3^- at high and low concentrations or NO_2^- at low concentration (high levels of NO_2^- result toxic for cells) were used as nitrogen sources for cell cultivation, RHA1::*nlpR* exhibited a delayed growth and lower cellular biomass production (principally at low-nitrogen levels) as compared to the WT strain (Fig. 4C-E). To analyze if growth retardation of RHA1::*nlpR* on MSM medium containing NO_3^- or NO_2^- was caused by the inability to reduce these compounds to ammonium, a Griess-Ilosvay assay was performed. After addition of the Griess-Ilosvay reagent to the supernatant of RHA1::*nlpR* strain grown overnight with NO_2^-, a more intense colour was observed in contrast to the lighter colour of the RHA1 strain. This result suggests the presence of remaining nitrite at this time point (Fig. 4F). Quantitative analyses of NO_2^- in the culture medium showed a slower consumption in the mutant in comparison with the WT strain (Fig. 4E). Similarly, the addition of zinc granules to the supernatant of the RHA1::*nlpR* grown overnight with NO_3^- and pre-treated with the Griess-Ilosvay reagent, promoted the appearance of a more intense colour in comparison with the lighter supernatant of the WT cell culture, demonstrating the presence of residual nitrate at this point (Fig. 4G). Altogether, these results demonstrate that although the *nlpR*

is not essential for growth with NO_3^- and NO_2^-, its disruption significantly affects the consumption rate of these compounds in RHA1.

The addition of cerulenin to the culture media also affected the dynamic of NO_3^-/NO_2^- reduction. A surplus of NO_2^- occurred in the supernatant of the RHA1 cells grown with NO_2^- plus cerulenin in comparison with cells growing without cerulenin, after addition of the Griess-Ilosvay reagent (Fig. 4F). Similar results were observed after addition of the Griess-Ilosvay reagent and zinc granules to cultures grown with NO_3^- plus cerulenin (Fig. 4G). These results demonstrated that cerulenin affected the uptake and assimilation of NO_3^-/NO_2^- to some extent, and correlated with the reduced expression of *nlpR* observed in cells growing in the presence of cerulenin (Fig. 3C).

Disruption of nlpR gene results in a decrease in TAG content and other lipid fractions in oleaginous rhodococcal strains under nitrogen-limiting conditions

Since a high expression of the *nlpR* gene is observed during TAG accumulation conditions and a low expression when the TAG biosynthesis is repressed in the presence of cerulenin, we speculated a functional relationship between this transcriptional regulator and lipogenesis in the oleaginous *R. jostii* RHA1. To explore this possibility, we analyzed the effect of *nlpR* disruption on TAG biosynthesis and accumulation under different growth conditions. No differences in TAG production occurred in RHA1::*nlpR* during cultivation of cells under nitrogen-rich conditions (LB medium or MSM medium supplemented with 1 g/L of NH_4Cl or KNO_3), as revealed by thin-layer chromatography (TLC) analysis (Fig. 5A). In contrast, the mutant strain exhibited a decrease in TAG content in comparison with the WT during cultivation under nitrogen-limiting conditions using different nitrogen sources (MSM0.1 with NH_4Cl, KNO_3 or $NaNO_2$) (Fig. 5A). This effect was even more noticeable when cells were cultivated in nitrogen-free medium (MSM0) as revealed by semiquantitative TLC analysis. In this case, a decrease of the other lipid fractions, such as monoacylglycerides (MAG), diacylglycerides (DAG) and free fatty acids (FFA), was also observed in RHA1::*nlpR* as compared to the RHA1 strain (Fig. 5B). Similar results were obtained during pulse-labeling experiments, resulting in a significant decrease in [^{14}C]-acetate incorporation into TAG fraction in RHA1::*nlpR* in comparison with the RHA1 cells (Fig. 5C). On the other hand, to expand this finding to other rhodococcal strains, we analyzed the effect of the disruption of *nlpR* gene in *R. opacus* PD630 and the indigenous strain *R. jostii* 602. Similarly to what it was observed

for the RHA1 mutant strain, disruption of *nlpR* gene in these two strains also resulted in a decrease in TAG accumulation and other lipid fractions in nitrogen-free media (Fig. 5D). A similar total fatty acid content (which is a good estimation of TAG content in oleaginous rhodococci) was observed in RHA1::*nlpR* and RHA1 during cultivation of cells under nitrogen-rich conditions (Table 1). In contrast, mutant strain showed a lower fatty acid content after cultivation of cells under nitrogen-limiting conditions, with a decrease of approximately 25% (w/w of the CDW) in the nitrogen-free medium (MSM0) (Table 1). To confirm that the observed phenotype was due to the *nlpR* gene disruption and to discard any polar effect associated to the disruption strategy used, the mutant strain was complemented with the vector pTip-QC2/*nlpR* and the resultant recombinant strain (RHA1::*nlpR*C) was cultured under nitrogen-free conditions to analyze the TAG production. Extrachromosomal *nlpR* expression restored the TAG content as well as the other lipid fractions at similar levels of the RHA1 cells (Fig. 5B and C). In accordance, total fatty acid content was higher in RHA1::*nlpR*C in comparison with RHA1::*nlpR* strain (Table 1). Fatty acid composition in RHA1::*nlpR* strain was the same to that reported previously for parental RHA1 strain (Hernández *et al.,* 2008) (data not shown). Finally, TLC of ^{14}C-labelled methyl esters of mycolic acids (MAMEs) and fatty acids (FAMEs) revealed a significant decrease of both fractions in RHA1::*nlpR* grown under nitrogen starvation conditions (Fig. 5E). Altogether, these results demonstrated that NlpR is contributing to the regulation of TAG biosynthesis as well as other lipid fractions somehow through an unknown mechanism in the oleaginous *R. jostii* RHA1 strain.

Overexpression of nlpR gene improves TAG accumulation in both oleaginous RHA1 and PD630 strains under nitrogen-rich conditions

Rhodococcal cells produce low amounts of TAG and high yields of cellular biomass during cultivation under nitrogen-rich conditions (Fig. 3A), when *nlpR* gene expression is low (Fig. 3B). In this context, we expected that overexpression of *nlpR* gene by addition of extrachromosomal copies, would promote a deregulation of TAG biosynthesis in oleaginous rhodococci under nitrogen-rich conditions. To confirm this hypothesis, we analyzed the effect of *nlpR* overexpression on TAG production in *R. jostii* RHA1 (RHA1 pTipQC2/*nlpR*) and *R. opacus* PD630 (PD630 pTipQC2/*nlpR*) during cultivation of cells in LB and MSM1 media. Lipid analyses confirmed that the induction of *nlpR* gene in RHA1 and PD630 strains cultivated in nitrogen-rich media produced

an increase of TAG as well as MAG, DAG and FFA fractions, without significant loss of cell biomass (Fig. 6). Additionally, TAG accumulation was improved by the addition of an exogenous free fatty acid (palmitic acid) to MSM1 medium plus glucose as carbon source, as is shown in Fig. 6. Quantitative analysis revealed an increase of 1.8-2-fold in total fatty acid content in RHA1 pTipQC2/*nlpR* and PD630 pTipQC2/*nlpR* strains grown in MSM1 medium in comparison with their respective control cells (RHA1 pTipQC2 and PD630 pTipQC2) (Table 1). Interestingly, overexpression of *nlpR* gene did not produce a further improvement in TAG content in RHA1 and PD630 strains during cultivation of cells under nitrogen-limiting conditions (MSM0.1 and MSM0 with gluconate and glucose as sole carbon sources) (data not shown). Altogether, these results suggested that the induction of *nlpR* gene activates the TAG accumulating-machinery to some extent, under conditions where it is usually turned-off.

The in silico analysis reveals the occurrence of a putative NlpR-binding site in the upstream regions of genes associated with nitrogen and lipid metabolisms in R. jostii RHA1

To analyze the occurrence of putative DNA binding sites for this protein in the upstream regions of genes of *R. jostii* RHA1, we used as reference the proposed DNA binding sequences for the NlpR ortholog in *Streptomyces coelicolor* (SCO2958, NnaR), which have been reported by Amin *et al.*, (2012). In this context, the upstream regions of *narK*, *nirB* and *nasA* genes of *S. coelicolor*, associated with nitrate/nitrite uptake and assimilation, exhibit an inverted repeated sequence proposed as the putative binding site for NnaR. This sequence is also conserved in the upstream regions of *nirB* and *narK* genes of *Mycobacterium tuberculosis*. Based on this background, we constructed a new curated Position Weight Matrix (PWM) by using all these sequences as well as the corresponding to the *narK* gene of RHA1 with both Regpredict web interface and MEME algorithm (Fig. 7A). The resulting motifs (Fig. 7B) were respectively submitted to the Regpredict platform and MAST server to scan the upstream regions of all the *R. jostii* RHA1 genes (coding region overlap was not accepted) in order to identify putative binding sites. In the case of Regpredict interface, the selected threshold score was based on that the consensus sequence should be conserved over 50% with respect to the highest score obtained for one of the sequences used to construct the PWM (MTU_NARK3, 8.06). As expected, and similarly to what it was observed with the SCO2958 protein in *S. coelicolor*, a putative binding site containing the inverted repeated motif

was found on the upstream region of *narK* (*RHA1_RS31125*) gene that is associated with nitrate/nitrite uptake in strain RHA1 (Fig. 7C). Although a binding site was no detected for the *nirD* (*RHA1_RS31135*) gene, a manual inspection of its intergenic region revealed a putative binding motif, which overlaps with the translation start site of the coding gene. Since no intergenic region is present between *narK* and *nirB* (*RHA1_RS31130*) or *nirD* and *nlpR* (*RHA1_RS31140*), this arrangement suggests a co-regulation of these genes (*narK-nirB* and *nirD-nlpR*) by the NlpR protein. On the other hand, another putative NlpR binding site was also found in the upstream region of *glnA1* (*RHA1_RS05590*) gene coding for a glutamine synthase. Interestingly, the putative NlpR binding sites occurring in upstream regions of all these genes associated with nitrogen metabolism were near to the binding site reported for GlnR, a well-known global regulator involved in control of genes associated with nitrogen metabolism in several actinobacteria (Wang *et al.*, 2015).

Among the putative NlpR binding sites identified by Regpredict interface appeared several sequences located in upstream regions of genes putatively involved in lipogenesis (Fig. 7C and Table 2). Some of these genes are putatively involved in fatty acid and TAG biosynthesis in *R. jostii* RHA1, such as the cluster (*RHA1_RS27565-RHA1_RS2755*) coding for a diacylglycerol acyltransferase (DGAT), a glycerol-3-phosphate acyltransferase (GPAT) and a putative bifunctional 1-acylglycerol 3-phosphate O-acyltransferase (AGPAT)/haloacid dehydrogenase HAD (with putative phosphatase activity), respectively. Similarly, a putative binding site was also found in upstream region of the pairs *RHA1_RS30955-RHA1_RS30960* (coding a Haloacid Dehalogenase-like Hydrolase and a DGAT, respectively) and *RHA1_RS02845-RHA1_RS02840* (coding a Putative lipase and a DGAT, respectively). Interestingly, most of these enzymes, putatively catalyze key reactions of the Kennedy pathway for TAG biosynthesis. Other identified motifs were present in upstream regions of genes related to fatty acid and TAG biosynthesis: *fas*I (*RHA1_RS06915*) coding for a fatty acid synthase type I, *plsC* (*RHA1_RS05380*) gene coding for a putative 1-acyl-sn-glycerol-3-phosphate O-acyltransferase (AGPAT), a *pap2* (*RHA1_RS00400*) gene coding for a phophatidic acid phosphatase type 2 and an *atf*3 (*RHA1_RS00220*) gene coding for another DGAT enzyme. In addition, a putative binding site was found in a region upstream of the *RHA1_RS05800-RHA1_RS05820* gene set, which are part of the FASII cluster as well as in the upstream region of the pair *RHA1_RS42560-RHA1_RS42555* coding for a biotin carboxyl carrier protein and an acetyl-

CoA carboxylase biotin carboxylase subunit, respectively. All these putative binding sites were also detected in upstream regions of the corresponding orthologous genes of *R. opacus* PD630 (data not shown). Another gene containing a putative NlpR binding site (with a score even higher than that observed for genes of nitrogen metabolism) was *panK*, which codes a pantothenate kinase enzyme involved in CoA metabolism. A complete list of putative binding sites on upstream regions of several genes is shown in Table 2.

Disruption of nlpR gene results in a decreased expression of genes associated with nitrogen and lipid metabolisms under nitrogen-limiting conditions

To confirm the role of NlpR in the expression of genes involved in nitrogen and lipid metabolisms, RT-PCR assays were performed from RNAs of RHA1 WT and the *nlpR* mutant strain, during cultivation of cells under nitrogen-limiting conditions (MSM0.1 with glucose as sole carbon source). For this experiment, we selected the following genes related to nitrogen and lipid metabolisms: *narK* gene (*RHA1_RS31125*) coding for the uptake system of nitrate/nitrite; *glnA1* gene (*RHA1_RS05590*) coding for a glutamine synthase; *fasI* (*RHA1_RS06915*) enconding the multifunctional fatty acid synthase; *plsC* (*RHA1_RS05380*) encoding a putative 1-acyl-sn-glycerol-3-phosphate O-acyltransferase (AGPAT) (Kennedy Pathway); *pap2* (*RHA1_RS00400*) gene coding for a phophatidic acid phosphatase type 2 (Kennedy Pathway); and *atf*3 (*RHA1_RS00220*) gene coding for a DGAT enzyme (Kennedy Pathway). We also included *mabR* (*RHA1_RS05800*) and *pccB* (*RHA1_RS05820*) genes encoding a putative transcriptional regulator (MabR) and a propionyl-CoA carboxylase (β subunit); respectively, both belonging to the FASII system. On the other hand, we used *fbpA* (*RHA1_RS19725*) gene coding for a putative mycolyltransferase as a negative reference gene, since no "NlpR-motif" was found in the upstream region of this gene with the settings parameters established. This mycolyltransferase is the ortholog of a diacylglycerol acyltransferase/mycolyltransferase Ag85A (Rv3804c) of *Mycobacterium tuberculosis*, which exhibited DGAT activity and an active contribution to TAG biosynthesis (Elamin *et al.*, 2011). As expected, the expression levels of *narK* decreased in the mutant strain in comparison with the WT strain (Fig.7D). On the other hand, *nlpR* mutant showed decreased transcription of *fasI*, *mabR* and *pccB* genes as compared to the WT strain (Fig.7D). Interestingly, transcripts of *plsC*, *pap*2 and *atf*3, which are part of the Kennedy pathway in strain RHA1, were not present in the mutant strain.

Finally, no significant changes in the transcription levels of *glnA1* and *fbpA* genes were observed between the WT and the *nlpR* mutant strain (Fig.7D). These results indicated that NlpR contributes to the modulation of the expression of genes involved in nitrogen metabolism as well as in fatty acid and TAG biosynthesis in the oleaginous *R. jostii* RHA1.

NlpR binds in vitro on the upstream regions of narK, nirD, fasI, plsC and atf3

To analyze if NlpR was able to bind on the upstream regions of genes related to nitrogen and lipid metabolisms in *R. jostii* RHA1, electrophoretic mobility shift assays (EMSAs) were performed. For this, purified His_6-NlpR protein from *E. coli* was used to analyze several DNA fragments corresponding to upstream regions of selected genes (*narK*, *nirD*, *fasI, plsC*, and *atf*3). An unspecific DNA fragment was used as negative control. EMSA assays showed a positive binding of His_6-NlpR to the upstream region of *narK*, *nirD*, *fasI*, *plsC* and *atf*3 genes at relative high protein concentration above 3.5 μM (data not shown). A complete binding was observed when using 8 μM of protein, whereas no binding was observed for negative control at this concentration (Fig. 7E).

Discussion

Oleaginous rhodococci may possess a regulatory network including transcriptional regulators controlling the expression of specific genes of TAG metabolism, and other global regulators which regulate this process at different hierarchical levels. These regulatory mechanisms occurring in oleaginous rhodococci are still unknown. In this study we found that a transcriptional regulator which is part of the regulatory network controlling the nitrogen metabolism also contributes to the modulation of TAG biosynthesis and accumulation in *R. jostii* RHA1 under nitrogen starvation conditions. For this reason, we named this transcriptional regulator as NlpR (Nitrogen lipid Regulator). Bioinformatic analyses revealed the occurrence of NlpR protein only in the actinobacteria group, including *Rhodococcus*, *Nocardia*, *Gordonia*, *Mycobacterium*, *Amycolatopsis* and *Streptomyces* genera, with the exception of members of *Corynebacterium* genus. Interestingly, the occurrence of NlpR seemed to be restricted to actinobacteria which possess the well-known global regulator of nitrogen metabolism, GlnR. This global regulator is also absent in corynebacteria. A comprehensive inspection of previous studies revealed that orthologs of *nlpR* are part of the GlnR regulon in related actinobacteria, such as *Streptomyces coelicolor* (Amin *et al.*, 2012), *Mycobacterium smegmatis* (Jeßberger *et al.*, 2013; Jenkins *et al.*, 2013), *M. tuberculosis*

(Williams *et al.*, 2015), *Amycolatopsis mediterranei* (Wang *et al.*, 2013) and *Saccharopolyspora erythraea* (Yao *et al.*, 2014).

NlpR-like proteins contain a conserved N-region highly similar to the uroporphyrinogen III synthase enzyme (HemD) and a C-region which resembles an OmpR-like domain associated with regulatory functions. The uroporphyrinogen III synthase enzymes have been associated with sirohaem biosynthesis, which is a prosthetic group of sulphite/nitrite reductases. Amin *et al.*, (2012) demonstrated that the HemD site at the N-region of the NnaR protein (Sco2958) does not possess enzymatic activity in *S. coelicolor*. Since *R. jostii* RHA1 and *R. opacus* PD630 possess additional proteins containing HemD regions putatively involved in sirohaem biosynthesis as cofactor for sulphite/nitrite reductases, this enzymatic reaction may not depend on the activity of NlpR. The N-terminal region of NlpR may interacts with a potential ligand such as a pyrrolic derivative, modulating the binding of the protein to DNA mediated by the C-terminal region. Similar mechanisms have been reported for some proteins known as trigger enzymes, in which their ability to modulate gene expression depend on ligand binding to specific domains (Commichau and Stülke, 2008; Commichau and Stülke, 2015).

In *R. jostii* RHA1 as well as in other *Rhodococcus* species, the *nlpR* gene is organized in a cluster together with *narK*, *nirB* and *nirD* genes, coding different components of the NO_3^-/NO_2^- uptake and reduction system (Iino *et al.*, 2012, Iino *et al.*, 2013). This conserved arrangement suggests a similar role of NlpR in strain RHA1 for the regulation of this cluster than their homologs NnaR in *S. coelicolor* and NasE in *Amycolatopsis mediterrenei* U32 (Amin *et al.*, 2012, Shao *et al.*, 2011). Results of this study indicated that *nlpR* gene of *R. jostii* RHA1 is induced during cultivation of cells under nitrogen deprivation conditions, which correlates well with previous studies. Dávila Costa *et al.*, (2015) reported a significant increase in NlpR, NarK, NirB, and NirD proteins abundance in *R. jostii* RHA1 under nitrogen starvation conditions in comparison with nitrogen-rich media. In other study, Iino *et al.*, (2012) demonstrated by transcriptomic analysis that the complete cluster where *nlpR* gene is located was highly expressed during incubation of *R. jostii* RHA1 cells in sterile soil containing both ammonium and nitrate at very low levels. Other transcriptomic data reported in different actinobacteria including *M. smegmatis* (Williams *et al.*, 2013), *M. tuberculosis* (Williams *et al.*, 2015), *R. opacus* PD630 (Chen *et al.*, 2015) and more recently, *R. jostii* RHA1 (Amara *et al.*, 2016), also revealed the upregulation of *nlpR* orthologs under ammonium-limiting

conditions. All these results confirmed that nitrogen deprivation is the key condition to induce the expression of NlpR in rhodococci and other actinobacteria, probably through a mechanism mediated by the global regulator GlnR, as has been reported for *S. coelicolor* (Amin *et al.*, 2012). Results of this study demonstrated that *nlpR* gene is not essential in *R. jostii* RHA1, although its disruption altered the normal growth of the cultures in NO_3^-/NO_2^- based media as well as in those media with limited concentrations of ammonium. Altogether, our study confirmed that NlpR is part of the regulatory network of rhodococci that controls nitrogen metabolism as response of cells to nitrogen deprivation.

Nitrogen starvation also triggers lipogenesis in oleaginous rhodococci (Alvarez and Steinbüchel, 2010). In this context, these microorganisms may possess some components which interconnect regulatory and metabolic networks for adapting their physiology during nutritional fluctuating conditions. The existence of some regulatory components that links nitrogen metabolism and lipogenesis in oleaginous rhodococci was considered in this study. The significant decrease of *nlpR* gene expression and the altered NO_3^-/NO_2^- assimilation by RHA1 cells after addition of cerulenin, suggested the possibility that NlpR could expand its regulatory function to lipogenesis. To explore this hypothesis, we generated a mutant strain for *nlpR* gene and analyzed its ability to produce several lipids in comparison with the WT strain of *R. jostii*. Interestingly, disruption of *nlpR* gene promoted a decrease of lipogenesis, including the production of TAG, DAG, MAG, and FFA fractions as compared to the WT strain. This effect was observed when cells were cultivated under nitrogen deprivation conditions but not in nitrogen-rich media. In addition, this effect was also observed in other taxonomically related strains, suggesting a common role of *nlpR* gene in oleaginous rhodococci. Furthermore, the *nlpR* complementation in the RHA1 mutant strain restored the TAG-accumulating phenotype of the WT strain, confirming its contribution into lipogenesis. On the other hand, the deregulation of *nlpR* gene by the effect of an inducible promoter produced a significant increase of lipid production in *R. jostii* RHA1 and *R. opacus* PD630, including TAG, DAG, MAG and FFA fractions, during cultivation of cells under nitrogen-rich conditions. This result suggested that the expression of NlpR activates the lipogenesis machinery in oleaginous rhodococci during an unfavorable condition for TAG accumulation. The addition of an exogenous fatty acid to the culture medium improved TAG production under nitrogen-rich conditions in *nlpR*-overexpressing rhodococcal cells. Besides its contribution to the regulation of TAG metabolism in

oleaginous rhodococci, overexpression experiments demonstrated that the manipulation of the transcriptional regulator NlpR involved in nitrogen metabolism and lipogenesis may be useful for the biotechnological production of oils and derivatives from nitrogen-rich industrial wastes.

To better understand the molecular basis of NlpR function related to the control of TAG metabolism, we performed some *in vivo* and *in vitro* studies using *R. jostii* RHA1 as microbial model. Based on putative binding sites recognized by the NlpR ortholog and EMSA assays reported for *S. coelicolor* (Amin *et al.*, 2012), we identified in this study several genes containing a putative "NlpR-motif" in their upstream sequences. Among them, we found several genes related to nitrogen and lipid metabolisms as well as some genes encoding putative enzymes of central metabolism and transcriptional regulators (Table 2). Similarly to *nlpR*, several of those genes are up-regulated under nitrogen-limiting conditions in RHA1 and PD630 strains (Table 2) (Amara *et al.*, 2016, Chen *et al.*, 2015). In this context, the NlpR protein may be considered as a pleiotropic regulator since it can influence the expression of genes of multiple pathways. A similar behavior have been reported for several other regulators in actinobacteria (Allenby *et al.,* 2012; Martin *et al.*, 2011; Martin *et al.*, 2012; Rabyk *et al.*, 2014; Liao *et al.*, 2015). In contrast to local specific regulators, global and pleiotropic regulators often possess more degenerate binding sites which provide them versatility in their binding affinity, as well as the possibility to control a higher number of genes. The occurrence of these non-canonical binding sites and its relationship with global regulation has been reported for several other transcriptional regulators in bacteria (Świątek-Połatyńska *et al.*, 2015).

The disruption of *nlpR* gene in *R. jostii* RHA1 resulted in a significant decrease of *narK* gene expression, which correlates with the altered growth of mutant cells with NO_3^-/NO_2^- as sole nitrogen sources, suggesting that NlpR is an activator for this gene, as has been reported for the NlpR ortholog (NnaR), which controls several genes associated with the NO_3^-/NO_2^- assimilation in *S. coelicolor* (Amin *et al.*, 2012). In addition, *nlpR* disruption also altered the expression of some genes involved in fatty acid and TAG biosynthesis, which correlates with the significant reduction of several lipids fractions in the oleaginous strain RHA1. In this context, the transcription of the *fasI* gene involved in fatty acid biosynthesis was significantly reduced by *nlpR* disruption in strain RHA1 (Fig. 7D) and may explains the decrease in the production of FFA, MAG and DAG fractions in addition to TAG, when disruptant cells are cultivated under nitrogen-limiting conditions. Moreover, the PCC (subunit β), which is part of the FASII system as well as the transcriptional

regulator MabR, which participates in the control of FASII expression in mycobacteria (Kurth *et al.*, 2009, Salzman *et al.*, 2010), seem to be also controlled by NlpR in *R. jostii* RHA1 to at least some extent (Table 2 and Fig. 7D). The lower expression of the respective genes (*pccB* and *mabR*) may explain the decrease of the mycolic acid fraction in mutant strain in addition to the fatty acid fraction (Fig. 5E). Finally, NlpR seemed to possess a strong influence on the expression of some genes involved in the Kennedy pathways for TAG biosynthesis in strain RHA1, since no transcripts of genes coding for an AGPAT (*pslC*), a PAP2 (*pap2*) and a DGAT (*atf*3) were detected after *nlpR* disruption (Fig. 7D). The involvement of this PAP2 enzyme and of an Atf3 ortholog in TAG biosynthesis and accumulation has been previously demonstrated in RHA1 and PD630 strains, respectively (Alvarez *et al.*, 2008; Hernández *et al.*, 2015).

Overexpression of NlpR in *R. jostii* RHA1 and *R. opacus* PD630 promoted a significant accumulation of TAG (up to 30% of CDW) under nitrogen-rich conditions, in which oleaginous rhodococci usually produce only low amounts of these lipids (Fig. 6 and Table 1). Altogether, these results indicate that NlpR mediates a positive transcriptional effect on lipogenesis processes and TAG accumulation in *R. jostii* RHA1 by the putative activation of the components of the TAG-accumulating machinery (or at least part of them) as it is shown in Fig. 8.

NlpR was able to bind on putative binding sites located upstream of *narK*, *nirD* from nitrogen metabolism, as well as *fasI*, *plsC*, and *atf*3 from lipid metabolism, as revealed *in vitro* studies (Fig 7E). However, the relative high concentration of NlpR used to achieve a complete binding in EMSA assays suggests the need of a yet unknown ligand or effector that could increase its affinity for the operator regions, as has been reported for other pleiotropic regulators (Tenconi *et al.*, 2015). The potential ligand, which may include NO_3^-/NO_2^- or metabolites from the synthesis of pyrrole- or lipid-derivatives, may bind to the domain located in the N-terminal region of the protein. To test this possibility, EMSA assays were performed using purified His-NlpR and the addition of $NaNO_2$, KNO_3, hydroxycobalamin, palmitoyl-CoA or palmitic acid, respectively. None of these molecules were able to improve the binding of NlpR to the different probes used (data not shown). Alternatively, a phosphorylation step or a correct protein interface interaction may be necessary for a correct dimerization and activation of the protein, increasing its affinity to DNA, as has been reported for other proteins of the OmpR-like family (Gao, *et al.*, 2010; Goyal *et al.*, 2011; Lin *et al.*, 2014). On the other hand, NlpR may also act as a co-activator of gene expression coordinately with

other more specific transcriptional regulators. In this context, the NlpR ortholog (NnaR) of *S. coelicolor* has been proposed as a co-activator of GlnR, which regulates the expression of several genes of nitrogen metabolism (Amin *et al.*, 2012). A similar NlpR-GlnR interaction may occur in *R. jostii* RHA1 for the activation of genes involved in nitrogen metabolism, since we found both, NlpR and GlnR-binding sites in the upstream sequences of genes of this metabolism. NlpR may also interact with other still unknown regulators or components to increase the *in vivo* DNA-binding affinity for the activation of lipogenesis genes. Thus, the complexity of these molecular mechanisms make unviable to reproduce *in vitro* the metabolic scenario and conditions occurring *in vivo*. Further studies are necessary to unravel the molecular basis of the NlpR-DNA interaction. Our results suggested that NlpR is a pleiotropic or global regulator participating in an intermediate hierarchical level probably in concert with other proteins and metabolic intermediates. This work provides a conceptual and experimental framework which can orientate research into the regulatory networks occurring in oleaginous rhodococci. Moreover, the identified transcriptional regulator may provide a new target for engineering oil production by rhodococci using nitrogen-rich industrial wastes.

Experimental procedures

Strains, plasmids and culture conditions

The strains and plasmids used in this work are listed in Table 3. *E. coli* strains were grown on solid or in liquid Luria-Bertani (LB) medium at 37 °C. *Rhodococcus* strains were cultivated aerobically at 30 °C in LB medium or minimal salt medium (MSM) according to Schlegel *et al.*, (1961). Gluconate or glucose were used in MSM media as sole carbon source at a final concentration of 1% (w/v). For nitrogen-limiting conditions to allow lipid accumulation, the concentration of ammonium chloride, sodium nitrite or potassium nitrate were reduced to 0.1 g/L (MSM0.1). In MSM0 culture medium, the addition of ammonium chloride or other nitrogen source was omitted. Cells were harvested at specific time-points, washed with a sterile NaCl solution (0.85%, w/v) and dried at 37 °C to constant weight for chemicals analyses. Antibiotics were used at the following final concentrations: 100 μg mL^{-1} ampicillin (Ap), 50 μg mL^{-1} kanamycin (Km) and 34 μg mL^{-1} chloramphenicol (Cm) in both *E. coli* and *Rhodococcus* strains. For overexpression analysis of genes under the thiostrepton promoter (P*tipA*) of pTip-QC2, a final concentration of 1-3 μg mL^{-1} of

thiostrepton was added to cell cultures at time zero. For inhibition of fatty acid synthesis, cerulenin (Sigma, St. Louis, MO) was used at 10 µg mL^{-1}. Cells were cultivated on LB medium at 30 °C for 24 h, washed, harvested, resuspended in MSM (MSM0 or MSM0.1 with sodium nitrite or potassium nitrate) medium containing sodium gluconate (1%, w/v) as the sole carbon plus cerulenin and incubated at 30 °C a specific time depending on the assay.

DNA analysis, amplification, cloning and sequencing

Chromosomal DNA, plasmids and DNA fragments were isolated and analyzed by standard methods. For specific DNA amplification of *nlpR* gene for cloning, disruption or heterologous expression as well as DNA probes amplification for EMSA assays, PCRs were performed with different specific primers listed in Table S1. The general thermocycler parameters used were as follows: 5 min at 94 °C, 30 cycles of 1 min at 94 °C, 30 seconds at 60 °C, 1 min at 72 °C, and finally 5 min at 72 °C. The PCR products were cloned into pGEM-T-easy vector and subjected to DNA sequencing.

Single nlpR gene disruption in rhodococci cells

To disrupt the *nlpR* gene, we used a derivate of pGEM-T-easy vector as a suicide plasmid and then followed a strategy previously reported for *R. erythropolis* SQ1 (Van der Geize *et al.*, 2000). For this, an internal 793-bp fragment of *RHA1_RS31140* (old locus tag RHA1_*ro06368*) gene (RHA1) was amplified by PCR using the primers ro06368F/ro06368R (Table S1) and then cloned in the pGEM-T-easy vector to get the plasmid pGEM-T-easy/"*nlpR"*. Then, a kanamycin resistant cassette obtained from the pUC4K plasmid was cloned into a *Pst*I restriction site of that plasmid to get pGEM-T-easy/"*nlpR*"-ΩKm (Fig. S3). The resulting suicide plasmid was transferred in *R. jostii* RHA1, *R. opacus* PD630 and *R. jostii* 602 by electroporation. After 5-6 days, recombination occurred and kanamycin resistant colonies were obtained, isolated and analyzed by colony-PCR using primers Ro06368R/Ro06368R (1 or 2) and M13F/Ro06368R (1 or 2) (Table S1). Positive kanamycin clones were re-confirmed by a second round of PCR but this time from purified chromosomal DNA.

Cloning of complete nlpR gene

Complete *nlpR* gene was amplified from total genomic DNA of *R. jostii* RHA1 by PCR using the primers Ro06368F and Ro06368R (Table S1). The resulting PCR product was cloned in pGEMT-easy vector, replicated in *E. coli* JM109 and verified by DNA sequencing. To achieve overexpression of *nlpR* in *Rhodococcus* strains, a *Nde*I/*Hin*dIII digest from pGEMT-easy/*nlpR* was subcloned into *Nde*I/*Hin*dIII site of the expression vector pTip-QC2 vector yielding pTip-QC2/*nlpR*. After confirming the correct reading frame via DNA sequencing, the resulting plasmid was transferred into RHA1 strain as well as into the PD630 strain (Gene Identity 99%) and selected on solid LB plates with chloramphenicol. The pTip-QC2/*nlpR* was also transferred into mutant RHA1 strain for complementation analysis and selected on solid LB plates with kanamycin plus chloramphenicol.

In order to heterologously express *nlpR* in *E. coli* BL21 (DE3), the *Nde*I/*Hin*dIII fragment from the pGEMT-easy/*nlpR* corresponding to the complete *nlpR* gene, was purified and subcloned into the *Nde*I/*Hin*dIII sites of pET28a expression vector (Novagen), yielding the plasmid pET28/*nlpR*. The pET28/*nlpR* plasmid was maintained in *E. coli* DH5α and transferred to *E. coli* BL21 (DE3) to perform recombinant protein (His_6-NlpR) purification by affinity chromatography. All plasmids described in this section are listed in Table 3.

DNA transfer and genotype screening in Rhodococcus cells

Suicide and replicative plasmids were transferred to *Rhodococcus* by electroporation. Electroporation assays were carried out as described by Kalscheuer *et al.,* (1999) using a Model 2510 electroporator (Eppendorf-Netheler-Hinz, Hamburg, Germany). The electrotransformants containing the complete *nlpR* gene under P*tip*A-promoter were checked using primers ThioF/Ro06368R1 or 2 (Table S1).

Heterologous expression of nlpR in E. coli and protein purification

E. coli BL21 (DE3) strain harboring plasmid pET28/*nlpR* was grown in LB at 37 °C until OD_{600nm} 0.6. Gene expression was induced by addition of 1 mM IPTG followed by overnight incubation at 30 °C and 180 rpm. Cells were harvested by centrifugation at 8,000 *g* for 20 min at 4 °C, washed twice with binding Buffer 20mM sodium phosphate, 500 mM NaCl, 20 mM imidazole, pH 7.4 and then resuspended in the same buffer. The next steps were all done at 4 °C. Cells disruption was

carried out with Fastprep 24 5-G homogenizer and lysis matrix B (MPBiomedicals, Ohio, USA.). The lysate was cleared by centrifugation at 13,000 *g* for 40 min to separate cell debris, and the supernatant was used to protein purification by affinity chromatography (His GraviTrap, GE Healthcare) according to manufacturer's instructions. Then, protein was concentrated and buffer interchanged with YM-30 Microcon® columns and resuspended in buffer (10mM Tris, 100 mM NaCl, 10% glycerol, pH 8). Purified His-NlpR protein was analyzed by SDS-PAGE and protein concentration was determined by Bradford assay using BSA as standard.

EMSA assays

DNA fragments containing the different upstream regions were amplified from *R. jostii* RHA1 genomic DNA by using Taq DNA polymerase (Promega) and primers listed in Table S1. The DNA fragments (10 ng) were mixed with different amounts of purified tagged protein (3.0 - 8.0 μM) and reaction buffer (10mM Tris pH 7.5, 100 mM NaCl, 1mM EDTA) in a total volume of 20 μL, which was incubated for 30 min at 30°C. After incubation, loading buffer was added and the fragments were separated on a 6% non-denaturing acrylamide gel. The gel was stained with SYBR Gold nucleic acid gel stain (Invitrogen). Unspecific DNA fragments were used as negative controls.

Total RNA isolation and RT-PCR assays

RNA was isolated from the 10 mL of cultured cells (OD_{600}: 2) as described by Gonçalves *et al.*, (2006) after 8 hours of incubation in MSM0 media. Prior to RNA isolation, cells were preserve by addition of 1/10 volume of 10% acid phenol (pH 5.0) in ethanol to the cell cultures. Cells were collected by centrifugation at 10,000 × g for 10 min at 4°C, suspended in 1.0 ml of ice-cold Tris-EDTA buffer (10mM Tris-HCl and 1 mM EDTA, pH 7.5) plus 2.0 ml RNA protect bacterial reagent (Roche), and incubated for 5 min at room temperature. The cells were then recovered by centrifugation at 13,000 × g for 2 min at room temperature, frozen, and stored at −80°C. Total RNA was isolated by vortexing with glass beads, treatment with hot phenol plus sodium dodecyl sulfate, removal of debris precipitated with acetate, phenol-chloroform extraction and isopropanol precipitation. RNA pellets were cleaned by using of a RNA isolation kit (Roche) and stored at −80°C until use. Alternatively, RNA was extracted from RHA1 and RHA1::*nlpR* strains grown in MSM0.1 using Direct Zol RNA MiniPrep (Zymo Research). For this, cells were resuspended in 1 mL of Quick zol and lysis was obtained by vortexing the cells with glass beads, sonication and

biruptor treatment (3 cycles of 7 minutes 30´´ ON-30´´ OFF) followed by incubation overnight at -80°C. The extraction protocol was followed according to manufacturer's instructions. Semiquantitative RT-PCR of *nlpR* gene was made by a one-step method using the primers Ro06368F and Ro06368R1 (Table S1) and Access RT-PCR kit according to manufacturer's instructions (Promega, USA, Madison). Semiquantitative RT-PCR of other genes was performed by a two-step method using second strand cDNA as template, generated with SuperScript III Reverse Transcriptase (Invitrogen) in presence of random hexamers. PCR was performed in a 25 µL reaction mixture containing 1 µL of a cDNA sample, 20 pmol of specific primers (Table S1), deoxynucleoside triphosphate (dNTP) mixture (0.25mM each), $MgCl_2$ 2 mM, Taq buffer 1x and 1 U of Taq Pegasus DNA polymerase (PB-L Prod BioL). RT-PCR cycling conditions were as follows: 94 ºC for 2 min followed by 28 cycles of 94 ºC for 30 s, 59 ºC for 30 s, and 72 ºC for 30 s. The amplified DNA was subjected to 2% agarose gel electrophoresis and visualized with ethidium bromide. As an internal control for each RT reaction 16S rRNA gene was included to check a constant RNA expression level between assayed samples.

Griess-Ilosvay assay

R. jostii RHA1 and the RHA1::*nlpR* mutant were grown at 30 °C overnight in LB medium. Cells were washed with NaCl solution (0.85%, w/v) and then a same inoculum was transferred into MSM medium containing KNO_3 or $NaNO_2$ at 0.1 g/L. Samples were collected at several time points or after incubation overnight. Cells were centrifuged at 10,000 *g*. A 300 µL volume of Griess-Ilosvay reagent (Merck) was added to 1 ml of cell supernatant, and accumulation of nitrite was determined by detection of a red colour.

Lipid analysis

Semi-quantitative analyses of total intracellular lipids in *Rhodococcus* cells were carried out by thin-layer chromatography (TLC). For this, 5 mg of dried cells was extracted with 300 µL chloroform/methanol (2:1, v/v) for 2 hours. 15 µL of chloroformic phase were subjected to TLC on silica Gel 60 F254 plates (Merck) using hexane/diethyl ether/acetic acid (80:20:1, v/v/v) as solvent for neutral lipids and free fatty acids analysis (Wältermann *et al.*, 2000). Lipid fractions were visualized by Iodine vapors staining.

In pulse labeling experiments for *de novo* lipid biosynthesis analysis, *Rhodococcus* was grown in 5 mL of MSM0.1 with glucose (1% w/v) as sole carbon source and labeled overnight with 3 μCi (neutral lipids) or 1 h with 1 μCi (FAMEs and MAMEs) of [^{14}C]acetate (50.5 mCi mmol^{-1}, Perkin-Elmer) at 30°C. Total neutral lipids were extracted from same cellular biomass and analyzed by TLC as described above. Alternatively, same c.p.m (50.000) of extracted lipids were loaded on TLC plate. Fatty acid methyl esters (FAMEs) and mycolic acid methyl esters (MAMEs) were obtained after treatment of the radiolabeled cell pellets containing the same number of cells as reported previously (Kurth *et al.*, 2009). The resulting solution of FAMEs and MAMEs was subjected to TLC and developed in hexane: ethyl acetate (9:1 v/v). The radioactivity incorporated into each lipid fraction was visualized by exposing the plates 2 days in Carestream® Kodak® BioMax® MR films or overnight to a storage screens capture (BAS-MS FujiFilm) followed of digitalization using a bio-imaging scanner system (Typhoon FLA-7000).

A colorimetric method (Duncombe, 1963; Wawrik, *et al.*, 2010) was performed for quantitative determination of total fatty acids in *Rhodococcus* cells. For this, dry cells (5-10 mg) were hydrolyzed with alkaline reagent (25% methanol in NaOH 1 N) at 95-100 °C for 3 h with vigorous agitation each 30 minutes. The soaps of fatty acids were neutralized with concentrated acetic acid and resultant free fatty acids were treated with the copper-reagent, extracted with chloroform and developed with the revealing reagent (diethyldithiocarbamate in 2-propanol). The resultant colored samples were spectrophotometrically read at 440 nm. The standard curve was performed with oleic acid as major representative fatty acid of TAG in rhodococcal cells. Alternatively, for qualitative and quantitative determination of total fatty acids, 5 mg of dry cells was subjected to methanolysis in the presence of 15 % (v/v) sulfuric acid as described by Brandl *et al.,* (1988), and the resulting acyl-methylesters were analyzed by gas chromatography (GC) using an HP 5890 A gas chromatograph equipped with an InnoWAX capillary column (30 m×0.53 mm×1 μm) and a flame ionization detector. The injection volume was 0.2 ml, and helium (13 mlmin−1) was used as a carrier gas. A temperature program was used for efficient separation of the methyl esters (90 °C for 5 min, temperature increase of 6 °C min−1, 220 °C for 10 min).

Bioinformatic analyses

To analyze *nlpR* and its homologs, we used all available bacterial genomes in NCBI database (http://www.ncbi.nlm.nih.gov/). Protein screening and alignments were carried out using BLAST 2.2.17 and CLUSTALW algorithms. Reference protein sequences were retrieved from the NCBI database. Identities were determined from alignments of full-length sequences. Phylogenetic analyses were carried out from curated sequences using the neighbor joining method with the program MEGA 5.1.

For finding putative binding sites of NlpR protein, computational approaches were used including the Reg predict platform (http://regpredict.lbl.gov/regpredict/) (Novichkov *et al.,* 2010), MEME and MAST tools (http://meme-suite.org/tools/meme; http://meme-suite.org/tools/mast).

Acknowledgments

This study was financially supported by the SCyT of the University of Patagonia San Juan Bosco, Project PIP2015-2017 Nro. 0529 (CONICET), Project PICT2012 Nro. 2031 (ANPCyT) and Project PICT2013 Nro. 0916, Argentina. Alvarez H.M., Hernández M.A, Gago G. and Gramajo H. are career investigators of the Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina.

References

Allenby, N.E.E., Laing, E., Bucca, G., Kierzek, A.M., Smith, C.P. (2012) Diverse control of metabolism and other cellular processes in *Streptomyces coelicolor* by the PhoP transcription factor: genome-wide identification of *in vivo* targets. *Nucleic Acids Research* **19**: 9543–9556.

Alvarez, A.F., Alvarez, H.M., Kalscheuer, R., Wältermann, M., Steinbüchel, A. (2008) Cloning and characterization of a gene involved in triacylglycerol biosynthesis and identification of additional homologous genes in the oleaginous bacterium *Rhodococcus opacus* PD630. *Microbiology* **154**:2327–2335.

Alvarez, H.M. (2010) Biotechnological production and significance of triacylglycerols and wax esters. In *Microbiology of hydrocarbons, oils, lipids, and derived compounds*. Kenneth, N.T. (ed). Heidelberg: Springer, pp. 2995–3002.

Alvarez, H.M. (2016) Triacylglycerol and wax ester-accumulating machinery in prokaryotes. *Biochimie* **120**:28–39.

Alvarez, H.M., Kalscheuer, R., Steinbüchel, A. (1997) Accumulation of storage lipids in species of *Rhodococcus* and *Nocardia* and effect of inhibitors and polyethylene glycol. *Fett-Lipid* **99**:239–246.

Alvarez, H.M., Mayer, F., Fabritius, D., Steinbüchel, A. (1996) Formation of intracytoplasmic lipid inclusion by *Rhodococcus opacus* PD630. *Arch Microbiol* **165**:377–386.

Alvarez, H.M., Steinbüchel, A. (2010) Physiology biochemistry and molecular biology of triacylglycerol accumulation by *Rhodococcus*. In *Biology of Rhodococcus*. Alvarez, H.M. (ed). Heidelberg: Microbiology monographs series. Springer, pp. 263–290.

Amara, S., Seghezzi, N., Otani, H., Diaz-Salazar, C., Liu, J., Eltis L.D. (2016) Characterization of key triacylglycerol biosynthesis processes in rhodococci. *Sci Rep* **6**:24985.

Amin, R., Reuther, J., Bera, A., Wohlleben, W., Mast, Y. (2012) A novel GlnR target gene, nnaR, is involved in nitrate/nitrite assimilation in *Streptomyces coelicolor*. *Microbiology* **158**(Pt 5):1172-82.

Arabolaza, A., D'Angelo, M., Comba, S., and Gramajo, H. (2010) FasR, a novel class of transcriptional regulator, governs the activation of fatty acid biosynthesis genes in *Streptomyces coelicolor*. Mol Microbiol **78:** 47–63.

Brandl, H., Gross, R.A., Lenz, R.W., Fuller, R.C. (1988) *Pseudomonas oleovorans* as a source of poly(β-hydroxyalkanoates) for potential applications as biodegradable polyesters. *Appl Environ Microbiol* **54**:1977–1982.

Chen, Y., Ding, Y., Yang, L., Yu, J., Liu, G., Wang, X., *et al.,* (2013) Integrated omics study delineates the dynamics of lipid droplets in *Rhodococcus opacus* PD630. *Nucleic Acids Res* **42**(2):1052–1064.

Commichau, F.M., Stülke, J. (2008) Trigger enzymes: bifunctional proteins active in metabolism and in controlling gene expression. *Mol Microbiol* **67**(4): 692–702.

Commichau, F.M., Stülke, J. (2015) Trigger Enzymes: Coordination of Metabolism and Virulence Gene Expression. *Microbiol Spectrum* **3**(4): MBP-0010-2014.

Dávila Costa, J.S., Herrero, O.M., Alvarez, H.M., Leichert L. (2015) Label-free and redox proteomic analyses of the triacylglycerol-accumulating *Rhodococcus jostii* RHA1. *Microbiology* **161**(Pt 3):593-610.

Ding,Y., Yang, L., Zhang, S.,Wang, Y., Du,Y., Pu, J., *et al.,* (2012) Identification of the major functional proteins of prokaryotic lipid droplets. *J Lipid Res* **53**:399–411.

Duncombe, W.G. (1963) The colorimetric micro-determination of long chain fatty acids. *Biochem J* **88**:7.

Elamin, A.A., Stehr, M., Spallek, R., Rohde, M., Singh, M. (2011) The *Mycobacterium tuberculosis* Ag85A is a novel diacylglycerol acyltransferase involved in lipid body formation. *Mol Microbiol* **81**(6):1577–1592.

Gao, R., and Stock, A.M. (2010) Molecular strategies for phosphorylation-mediated regulation of response regulator activity. *Curr Opin Microbiol* **13**: 160–167.

Gonçalves, E.R., Hara, H., Miyazawa, D., Davies, J.E., Eltis, L.D., Mohn, W.W. (2006) Transcriptomic assessment of isozymes in the biphenyl pathway of *Rhodococcus* sp. strain RHA1. *Appl Environ Microbiol* **72**(9):6183–6193.

Goyal, R., Das, A.K., Singh, R., Singh, P.K., Korpole, S., Sarkar, D. (2011) Phosphorylation of PhoP protein plays direct regulatory role in lipid biosynthesis of *Mycobacterium tuberculosis*. *J Biol Chem* **286**:45197–45208.

Hanahan, D. (1983) Studies on transformation of *Escherichia coli* with plasmids. *J Mol Biol* **166**:557–580.

Hernández, M.A., Alvarez, H.M. (2010) Glycogen formation by *Rhodococcus* species and the effect of inhibition of lipid biosynthesis on glycogen accumulation in *Rhodococcus opacus* PD630. *FEMS Microbiol Lett* **312**:93–99.

Hernández, M.A., Arabolaza, A., Rodríguez, E., Gramajo, H., Alvarez, H.M. (2013) The *atf*2 gene is involved in triacylglycerol biosynthesis and accumulation in the oleaginous *Rhodococcus opacus* PD630. *Appl Microbiol Biotechnol* **97**:2119–2130.

Hernández, M.A., Comba, S., Arabolaza, A., Gramajo, H., Alvarez (2015) Overexpression of a phosphatidic acid phosphatase type 2 leads to an increase in triacylglycerol production in oleaginous *Rhodococcus* strains. *Appl Microbiol Biotechnol* **99**(5):219–1207.

Hernández, M.A., Mohn, W.W., Martínez, E., Rost, E., Alvarez, A.F., Alvarez, H.M. (2008) Biosynthesis of storage compounds by *Rhodococcus jostii* RHA1 and global identification of genes involved in their metabolism. *BMC Genomics* **12**:600

Holder, J.W., Ulrich, J.C., DeBono, A.C., Godfrey, P.A., Desjardins, C.A., Zucker, J., *et al.,* (2011) Comparative and functional genomics of *Rhodococcus opacus* PD630 for biofuels development. *PLoS Genet* **7**(9):e1002219. doi: 10.1371/journal.pgen.1002219.

Iino, T., Miyauchi, K., Kasai, D., Masai, E., Fukuda, M. (2013) Characterization of nitrate and nitrite utilization system in *Rhodococcus jostii* RHA1. *J Biosci Bioeng* **115**(6):600–6.

Iino, T., Wang, Y., Miyauchi, K., Kasai, D., Masai, E., Fujii, T., Ogawa, N., Fukuda, M. (2012) Specific Gene Responses of *Rhodococcus jostii* RHA1 during Growth in Soil. *Appl Environ Microbiol* **78**(19):6954–62.

Jenkins, V.A., Barton, G.R., Robertson, B.D., Williams, K.J. (2013) Genome wide analysis of the complete GlnR nitrogen-response regulon in *Mycobacterium smegmatis*. *BMC Genomics* **14**:301.

Jeßberger, N., Lu, Y., Amon, J., Titgemeyer, F., Sonnewald, S., Reid, S., Burkovski, A. (2013) Nitrogen starvation-induced transcriptome alterations and influence of transcription regulator mutants in *Mycobacterium smegmatis*. *BMC Research Notes* **6**:482.

Kalscheuer, R., Arensk̈otter, M., Steinbüchel, A. (1999) Establishment of a gene transfer system for *Rhodococcus opacus* PD630 based on electroporation and its application for recombinant biosynthesis of poly(3-hydroxyalkanoic acids). *Appl Microbiol Biotechnol* **52**:508–515.

Kurth, D.G., Gago, G.M., de la Iglesia, A., Bazet, L.B., Lin, T.W., Morbidoni, H.R., *et al.,* (2009) Accase 6 is the essential acetyl-CoA carboxylase involved in fatty acid and mycolic acid biosynthesis in mycobacteria. *Microbiology* **155**:2664–2675.

Liao, C.H, Xu, Y., Rigali, S., Ye, B.C. (2015) DasR is a pleiotropic regulator required for antibiotic production, pigment biosynthesis, and morphological development in *Saccharopolyspora erythraea*. *Appl Microbiol Biotechnol* **99**(23):10215–24.

Liao, C.H., Yao, L., Xu,Y., Liu, W.B., Zhou,Y., Ye, B.C. (2015) Nitrogen regulator GlnR controls uptake and utilization of non-phosphotransferase-system carbon sources in actinomycetes. *Proc Natl Acad Sci* **112**(51):15630–5.

Lin, W., Wang, Y., Han, X., Zhang, Z., Wang, C., Wang, J., *et al.,* (2014) Atypical OmpR/PhoB subfamily response regulator GlnR of actinomycetes functions as a homodimer, stabilized by the unphosphorylated conserved Asp-focused charge interactions. *J Biol Chem* **289**(22):15413–25.

MacEachran, D.P., Prophete, M.E., Sinskey, A.J. (2010) The *Rhodococcus opacus* PD630 heparin-binding hemagglutinin homolog TadA mediates lipid body formation. *Appl Environ Microbiol* **76**(21):7217–25.

MacEachran, D.P., Sinskey, A.J. (2013) The *Rhodococcus opacus* TadD protein mediates triacylglycerol metabolism by regulating intracellular NAD(P)H pools. *Microb Cell Fact* **12**:104.

Martín, J.F., Liras, P. (2012) Cascades and networks of regulatory genes that control antibiotic biosynthesis. *Subcell Biochem* **64**:115–38.

Martín, J.F., Sola-Landa, A., Santos-Beneit, F., Fernández-Martínez, L.T., Prieto, C., Rodríguez-García, A. (2011) Cross-talk of global nutritional regulators in the control of primary and secondary metabolism in *Streptomyces*. *Microb Biotechnol* **4**(2):165–74.

Nakashima, N., Tamura, T. (2004) Isolation and characterization of a rolling-circle-type plasmid from *Rhodococcus erythropolis* and application of the plasmid to multiple recombinant-protein expression. *Appl Environ Microbiol* **70**:5557–5568.

Novichkov, P.S., Rodionov, D.A., Stavrovskaya, E.D., Novichkova, E.S., Kazakov, A.E., *et al.,* (2010) RegPredict. An integrated system for regulon inference in prokaryotes by comparative genomics approach. *Nucleic Acids Res* **38**:W299–W307.

Rabyk, M.V., Ostash, B.O., Fedorenko, V.O. (2014) Gene networks that regulate secondary metabolism in actinomycetes: pleiotropic regulators. *Tsitol Genet* **48**(1):67–82.

Salzman, V., Mondino, S., Sala, C., Cole, S.T., Gago, G., Gramajo, H. (2010) Transcriptional regulation of lipid homeostasis in mycobacteria. *Mol Microbiol* **78**(1):64–77.

Schlegel, H.G., Kaltwasser, H., Gottschalk, G. (1961) Ein Submersverfahren zur Kultur Wasserstoff oxydierender Bakterien: Wachstumsphysiologische Untersuchungen. *Arch Mikrobiol* **38**:209–222.

Seto, M., Kimbara, K., Shimura, M., Hatta, T., Fukuda, M., Yano, K. (1995) A novel transformation of polychlorinated biphenyls by *Rhodococcus* sp. strain RHA1. *Appl Environ Microbiol* **61**(9):3353–3358.

Shao, Z., Gao, J., Ding, X., Wang, J., Chiao, J., Zhao, G. (2011) Identification and functional analysis of a nitrate assimilation operon *nasACKBDEF* from *Amycolatopsis mediterranei* U32. *Arch Microbiol* **193**(7):463–477.

Silva, R.A., Grossi, V., Olivera, N.L., Alvarez, H.M. (2010) Characterization of indigenous *Rhodococcus* sp. 602, a strain able to accumulate triacylglycerides from naphthyl compounds under nitrogen-starved conditions. *Res Microbiol* **161**(3):198–207.

Stamford, N.P., Capretta, A. Battersby, A.R. (1995) Expression, purification and characterization of the product from the *Bacillus subtilis hemD* gene, uroporphyrinogen III synthase. *Eur J Biochem* **231**:236–241.

Świątek-Połatyńska, M.A., Bucca, G., Laing, E., Gubbens, J., Titgemeyer, F., Smith, C.P., Rigali, S., van Wezel, G.P. (2015) Genome-wide analysis of *in vivo* binding of the master regulator DasR in *Streptomyces coelicolor* identifies novel non-canonical targets. *PLoS ONE* **10**(4): e0122479. doi:10.1371/journal.pone.0122479

Tenconi, E., Urem, M., Świątek-Połatyńska, M.A., Titgemeyer, F., Muller, Y.A., van Wezel, G.P., Rigali, S. (2015) Multiple allosteric effectors control the affinity of DasR for its target sites. *Biochem Biophys Res Commun* **464**(1):324–9.

Van der Geize, R., Hessels, G.I., Van Gerwen, R., Vrijbloed, J.W., Van der Meijden, P., Dijkhuizen, L. (2000) Targeted disruption of the *kstD* gene encoding a 3-ketosteroid Δ1-dehydrogenase isoenzyme of *Rhodococcus erythropolis* strain SQ1. *Appl Environ Microbiol* **66**:2029–2036.

Villalba, M.S., Alvarez, H.M. (2014) Identification of a novel ATP-binding cassette transporter involved in long-chain fatty acid import and its role in triacylglycerol accumulation in *Rhodococcus jostii* RHA1. *Microbiology* **16**:1523-1532.

Wältermann, M., Luftmann, H., Baumeister, D., Kalscheuer, R., Steinbüchel, A. (2000) *Rhodococcus opacus* PD630 as a source of high-value single cell oil? Isolation and characterization of triacylgycerols and other storage lipids. *Microbiology* 146:1143–1149.

Wang, J., Wang, Y., Zhao, G.P. (2015) Precise characterization of GlnR Box in actinomycetes. *Biochem Biophys Res Commun* **458**(3):605–607.

Wang, Y., Wang, J.Z., Shao, Z.H., Yuan, H., Lu, Y.H., Jiang, W.H., *et al.,* (2013) Three of four GlnR binding sites are essential for GlnR-mediated activation of transcription of the *Amycolatopsis mediterranei nas* operon. *J Bacteriol* **195**(11):2595–25602.

Wawrik, B., Harriman, B.H. (2010) Rapid, colorimetric quantification of lipid from algal cultures. J *Microbiol Methods* **80**:262–266.

Williams, K.J., Bryant, W.A., Jenkins, V.A., Barton, G.R., Witney, A.A., Pinney, J.W., Robertson, B.D. (2013) Deciphering the response of *Mycobacterium smegmatis* to nitrogen stress using bipartite active modules. *BMC Genomics* **14**:436.

Williams, K.J., Jenkins, V.A., Barton, G.R., Bryant, W.A., Krishnan, N., Robertson, B.D. (2015) Deciphering the metabolic response of *Mycobacterium tuberculosis* to nitrogen stress. *Mol Microbiol* **97**(6):1142–1157.

Yao, L.L., Liao, C.H., Huang, G. Zhou, Y. Rigali, S. Zhang, B., Ye, B.C. (2014) GlnR-mediated regulation of nitrogen metabolism in the actinomycete *Saccharopolyspora erythraea*. *Appl Microbiol Biotechnol* **98**(18):7935-48.

Author Contributions

MAH carried out the experimental and bioinformatic studies; participated in the sequence analyses and alignments, in the design of the study and drafted the manuscript. JL and GG participated in the experimental studies on RT-PCR and EMSA and in the interpretation and discussion of data. HG participated in the coordination of the study, in the interpretation of data and helped to draft the manuscript. HMA conceived the study and participated in its design and coordination, interpretation of data, and helped to draft manuscript. All authors read and approved the final manuscript.

Conflict of Interest

None declared.

Figures legends

Figure 1: Occurrence of NlpR in Actinobacteria.

Multiple sequence alignment of full NlpR protein of RHA1 strain and orthologues in other actinobacteria. Identical amino acid residues are shaded in black and similar amino acid residues appear in light grey. The trans_reg_C region containing a helix-turn-helix domain is boxed and putative DNA binding sites are highlighted with asterisks.

Figure 2: Genomic context of *nlpR* gene in several actinobacteria.

Arrows indicate positions, relative length and transcriptional orientation. Ref: *nirB* (*nasB*) and *nirD* (*nasD*) represents genes for nitrite reductase large and small subunits, respectively; *narK* (*nasK*) and *narK3*, nitrate/nitrite uptake genes; *nasA* and *nasC*, genes for nitrate reductase catalytic and electron transfer subunits, respectively.

Figure 3: TAG accumulation and *nlpR* expression analysis.

A. TAG accumulation analyzed by TLC under rich-nitrogen (MSM1) and free-nitrogen conditions (MSM0) using RHA1 strain as model. B. Expression of *nlpR* gene analysis in both RHA1 and PD30 strains under rich-nitrogen (MSM1) and free-nitrogen conditions (MSM0). C. Expression of *nlpR* gene analysis in both RHA1 and PD30 strains under free-nitrogen conditions (MSM0) and free-nitrogen conditions plus cerulenin (MSM0+cerulenin).

Figure 4: Effect of *nlpR* disruption on growth with several nitrogen sources and Griess-Ilosvay assay.

Growth profiles of wild type strain (RHA1) and mutant strain (RHA1::*nlpR*) grown in MSM minimal medium supplemented with: A-B. Ammonium chloride, C-D. Potassium nitrate, E. Sodium nitrite. F. Determination of nitrite with Griess-Ilosvay reagent in supernatants of wild type strain (RHA1), mutant strain (RHA1::*nlpR*) and wild type strain with cerulenin (RHA1+cerulenin) grown overnight with sodium nitrite at 0.1 g/L. G. Determination of nitrite with Griess-Ilosvay reagent and Zn granules in supernatants of wild type strain (RHA1), mutant strain (RHA1::*nlpR*) and wild type strain with cerulenin (RHA1+cerulenin) grown overnight with potassium nitrate 0.1 g/L.

Figure 5. Effect of *nlpR* disruption in TAG accumulation and other lipid fractions in oleaginous rhodococci under nitrogen-limiting conditions.

A. TAG fraction analysis by TLC in wild type RHA1 strain (RHA1) and mutant RHA1 strain (RHA1::*nlpR*) grown in LB or MSM media with different nitrogen sources at several concentration levels. Cells were grown in LB medium for 24 h, harvested, washed and then incubated in LB, MSM1 or MSM0.1 for 48 hours with gluconate (1% w/v) as sole carbon source. B. Total lipid analysis by TLC in RHA1, RHA1::*nlpR* and complemented strain (RHA1::*nlpR* C). Cells were grown in LB medium for 24 h, harvested, washed and then incubated in MSM0 for 48 hours with gluconate (1% w/v) as sole carbon source. C. Pulse-labeling experiment for analysis of *de novo* lipid biosynthesis in RHA1, RHA1::*nlpR* and RHA1::*nlpR* C. Cells were grown in LB medium for 24 h, harvested, washed and then incubated overnight in MSM0.1 with glucose (1% w/v) as sole carbon source and [^{14}C]-acetate. D. Total lipid analysis by TLC in wild type strains (PD630 and 602) and mutant strains (PD630::*nlpR* and 602::*nlpR*). Cells were grown as in B. E. Thin-layer chromatography (TLC) of ^{14}C-labelled methyl esters of mycolic acids (MAMEs) and FA (FAMEs) extracted from RHA1 and RHA1::*nlpR* strains grown under nitrogen-limiting conditions. Ref.: TAG, triacylglycerol; FFA, free fatty acid; DAG, diacylglycerol; MAG, monoacylglycerol; PL, phospholipid.

Figure 6: Effect of *nlpR* overexpression in TAG accumulation and other lipid fractions in oleaginous rhodococci under nitrogen-rich conditions.

Overexpressing *nlpR*-recombinant strains (RHA1 pTipQC2/*nlpR* and PD630 pTipQC2/*nlpR*) and control cells (RHA1 pTipQC2 and PD630 pTipQC2) were grown in LB medium for 24 h, harvested, washed, and then inoculated in LB or MSM1 medium with gluconate as carbon source and incubated for 48h. MSM1 media containing palmitic acid (0.1% w/v) are highlighted with an asterisk. TAG triacylglycerol, FFA free fatty acid, DAG diacylglycerol, MAG monoacylglycerol, PL phospholipid.

Figure 7: NlpR modulates genes associated with nitrogen and lipid metabolisms in *R. jostii* RHA1.

A. Upstream regions of different nitrate/nitrite genes from *S. coelicolor* (SCO), *M. tuberculosis*

(MTU) and *Rhodococcus* (RHO) used for Position weight matrix (PWM) construction. Grey boxes indicate the predicted NlpR recognition sites. B. Putative binding site LOGO representation of the consensus DNA-binding motif identified for NlpR by Regpredict interface (left) and MEME tool (right). C. Alignment of putative *cis*-elements found in upstream regions of genes of nitrogen and lipid metabolism. D. RT-PCR analyses of genes putatively involved in nitrogen metabolism (*narK* and *glnA1*) and lipid metabolism (*plsC*, *fasI*, *pap2*, *atf*3, *mabR*, *pccB*, *fbpA*) in RHA1 and RHA1::*nlpR* under low-nitrogen conditions. Gene expression levels of these genes were compared and normalized on the basis of copy numbers of 16S-ribosomal transcripts (*16S*). A negative control (total RNA) was also used to discard amplification from genomic contamination. E. EMSA assay using purified His-NlpR and the *narK*, *nirD fasI*, *plsC* and *atf*3 upstream regions. Purified His-NlpR at 8.0 µM was incubated with 10 ng of DNA probes. Unspecific DNA probe was used as negative control. Shift DNA probes are shown by arrows.

Figure 8: Summary of some aspects of the nitrogen and lipid metabolism in oleaginous rhodococci and a model of NlpR regulation. The NO_3^-/NO_2^- uptake and assimilation systems as well as some reactions involved in lipogenesis are indicated. A model for regulation of NO_3^-/NO_2^- uptake and assimilation and lipogenesis by NlpR and its interplay with nitrogen deprivation conditions is proposed. Continuous lines represent a regulatory effect by NlpR based on experimental data obtained in this study, whereas dotted lines represent a putative regulatory effect by NlpR.

Table 1: Total fatty acid content in recombinant strains in several culture media after 48h

Strain	Culture medium	Carbon source	% Total fatty acid content (CDW)
RHA1	LB	-	11.0
RHA1::*nlpR*	LB	-	9.8
RHA1	MSM1 (NH_4Cl 1g/L)	Gluconate	16.1
RHA1::*nlpR*	MSM1 (NH_4Cl 1g/L)	Gluconate	14.0
RHA1	MSM1 (KNO_3 1g/L)	Gluconate	15.8
RHA1::*nlpR*	MSM1 (KNO_3 1g/L)	Gluconate	12.0
RHA1	MSM0.1 (NH_4Cl 0.1g/L)	Gluconate	51.4
RHA1::*nlpR*	MSM0.1 (NH_4Cl 0.1g/L)	Gluconate	30.4
RHA1	MSM0.1 (KNO_3 0.1g/L)	Gluconate	34.2
RHA1::*nlpR*	MSM0.1 (KNO_3 0.1g/L)	Gluconate	20.9
RHA1	MSM0.1 ($NaNO_2$ 0.1g/L)	Gluconate	39.2
RHA1::*nlpR*	MSM0.1 ($NaNO_2$ 0.1g/L)	Gluconate	28.6
RHA1	MSM0	Gluconate	54.3
RHA1::*nlpR*	MSM0	Gluconate	31.3
RHA1::*nlpR*C	MSM0	Gluconate	46.6
RHA1 pTipQC2	MSM1 (NH_4Cl 1g/L)	Gluconate	16.8
RHA1 pTipQC2/*nlpR*	MSM1 (NH_4Cl 1g/L)	Gluconate	29.6
PD630 pTipQC2	MSM1 (NH_4Cl 1g/L)	Gluconate	17.1
PD630 pTipQC2/*nlpR*	MSM1 (NH_4Cl 1g/L)	Gluconate	35.0

Table 2: Occurrence of putative NlpR binding sites identified by Regpredict interface on upstream regions of genes associated mainly with nitrogen and lipid metabolism in RHA1 strain

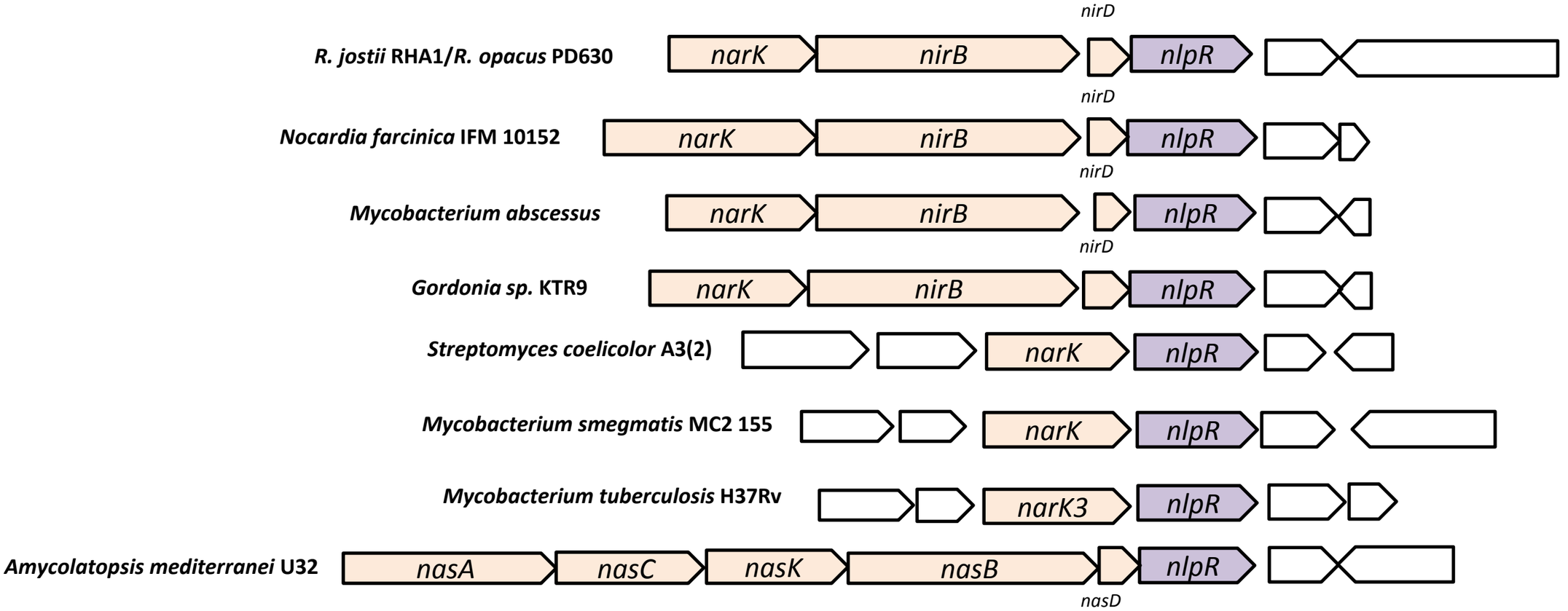
Gene ID	Gen name	Product(s)	Position	Motif	Threeshold score
RHA1_RS28600	*panK*	Pantothenate kinase (*) (‡)	-36	CTCACTC-(14)-GTGTGAG	7.55
RHA1_RS34180- ***RHA1_RS34185***	*-fabG*	Transcriptional regulator-3 ketoacyl-ACP reductase (FabG) (*)	-41	GTCACAG-(13)-GTGTGAG	7.20
RHA1_RS05590	*glnA1*	Glutamine synthetase* (†)(‡)	-178	CTCACAC-(14)-GTGTGCG	7.04
RHA1_RS31125-RHA1_RS31140	*narK-nlpR*	Major facilitator transporter-Nitrite reductase large and small subunit-NlpR (*) (†)(‡)	-115	CTCACAT-(13)-AAGTGAG	6.85
RHA1_RS29320		Malate decarboxilase (oxalacetate decarboxilating) NADP-dependent* (†)(‡)	-318	CTCACAG-(14)-ACGTGAG	6.61
RHA1_RS27565-RHA1_RS2755		DGAT-glycerol-3-phosphate acyltransferase-HAD hydrolase (*)	-121	CTCACCT-(13)-CGGTGCG	6.12
RHA1_RS10890		ABC aminoacid transporter*	-253	CGCTCAC-(13)-CGGTGAG	6.02
RHA1_RS39800		Malate decarboxilase (oxalacetate decarboxilating) NADP-dependent* (†)(‡)	-221	ATCCCCG-(14)-CGGTGAG	6.01
RHA1_RS00400		Phosphatidic acid phosphatase type2 (PAP2) (*)(‡)	-39	CTCAGCG_(13)-CGGTGAC	5.98
RHA1_RS00395		DNA-binding response regulator (*)(†)	-92	GTCACCG-(13)-CGCTGAG	5.98
RHA1_RS10105		Long-chain-fatty-acid-CoA ligase (*)	-84	GTCACAG-(13)-GTCTGAT	5.93
RHA1_RS40505		Lipase (†)	-353	CTCACTC-(13)-GTGGGGG	5.91
RHA1_RS39865		Enoyl coA hydratase (*)	-273	CTCACAG-(14)-CGCCGAG	5.80
RHA1_RS22410		Long-chain-fatty-acid-CoA ligase (*)	-157	CTTTCAG-(14)-CGGTGAG	5.78
RHA1_RS03595		PHA Synthase	-211	CTCATCC-(14)-GAGGGAG	5.78
RHA1_RS10030		Haloacid Dehalogenase-like Hydrolase	-39	CTCGCAC-(13)-CGGCGAG	5.73
RHA1_RS23045		TAG lipase (*)	-62	CTCATCT-(14)-AAGTGAG	5.70
RHA1_RS00975		Acyl-CoA synthetase	-35	GTCACAG-(14)-CTCGGAG	5.69
RHA1_RS20885		Lipase (*)	-78	CTCCCCC-(14)-GTGTGCT	5.64
RHA1_RS25160		Esterase	-161	CACCCAC-(14)-CGGTGAC	5.40
RHA1_RS42570		Acyl-CoA synthetase	-29	CTCAGCC-(13)-CTCTGCG	5.34
RHA1_RS30955-RHA1_RS30960	*-atf*10	Haloacid Dehalogenase-like Hydrolase-DGAT (†)	-104	CTTACCC-(13)-ATGAGAC	5.29
RHA1_RS00220	*atf*3	DGAT	-119	CGCCCAT-(13)-ATGTGAC	5.10
RHA1_RS06915	*fasI*	Fatty acid synthase (FASI) (†)(‡)	-100	GACACCC-(13)-ACGTGAG	5.06
RHA1_RS02845-RHA1_RS02840	*-atf*5	Putative lipase-DGAT	-102	CTGAGAC-(13)-GCGAGAG	4.92
ro04106		possible lipase/acylhydrolase (†)(‡)	-32	CTCCGAG-(13)-TGGGGAG	4.91
ro02708		Long-chain-fatty-acid-CoA ligase (†)	-313	CGCACAT-(14)-AAGTGGG	4.81
RHA1_RS05380	*plsC*	1-acyl-sn-glycerol-3-phosphate acyltransferase (†)	-123	CCCACTG-(13)-GTGTGAT	4.76
RHA1_RS05800-RHA1_RS05820	*mabR-pccB*	Cluster FASII (ro01199-ro01202†)	-61	CCCACCG-(14)-GTGGGAC	4.75
RHA1_RS19965		Lipase	-96	GTCAGAG-(14)-GGGAAAG	4.75
RHA1_RS42560-RHA1_RS42555	*-accC2*	Biotin carboxyl carrier protein-Acetyl-coA carboxilase biotin carboxilase subunit (†)	-78	CTCGCTG-(13)-CGGGGAC	4.70

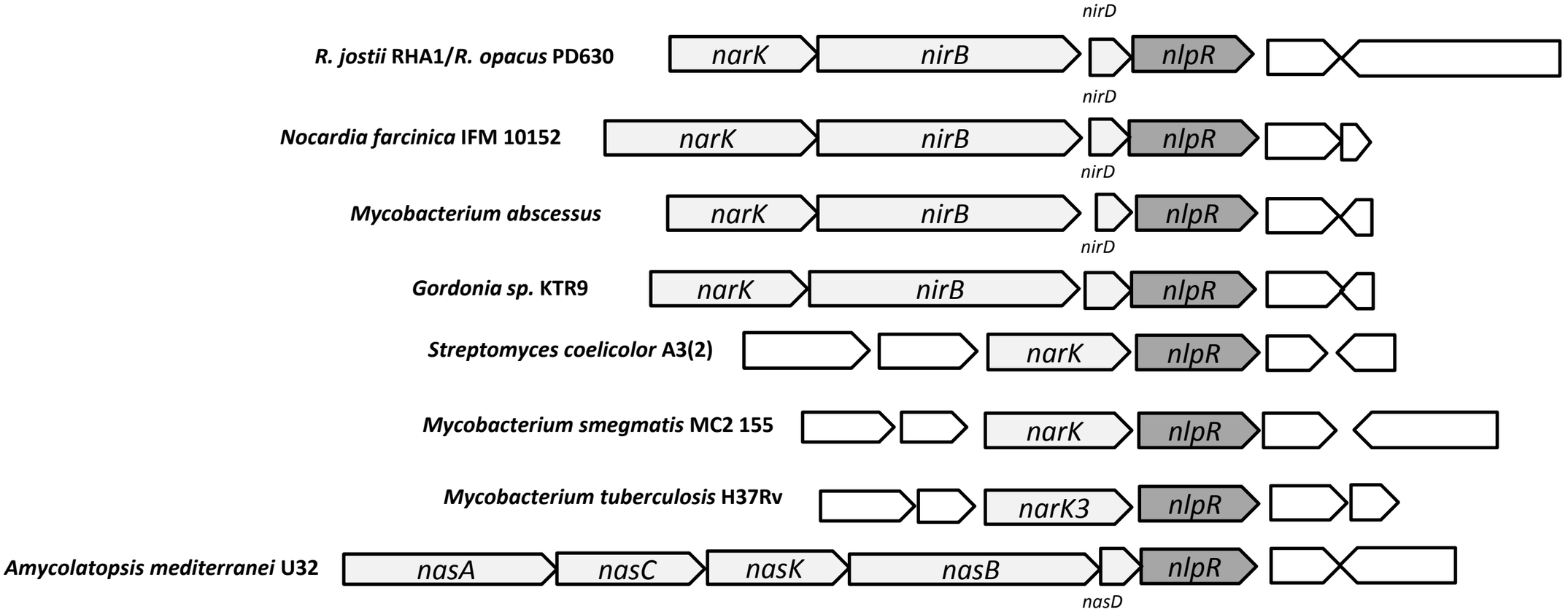
RHA1_RS03600		Poly(3-hydroxyalkanoate) depolymerase	-299	ACCACAC-(13)-GGCCGAG	4.65
RHA1_RS39870		Long-chain-fatty-acid-CoA ligase(*)(†)	-193	CTCTCGC-(14)-TGTTGAG	4.58
RHA1_RS22415	*fadE*	acyl-CoA dehydrogenase (*)	-200	ACCACCC-(13)-GGGCGAC	4.52
RHA1_RS20065		Esterase	-99	AGCACAC-(14)-CTCTAAG	4.51
RHA1_RS13005		TAG lipase (†)	-84	CGCCCCC-(13)-GGGTGCT	4.46
RHA1_RS31855		Enoyl-coA hydratase (†)(‡)	-123	CGCCCCC-(14)-GGGGGCG	4.46
RHA1_RS13165		Long-chain-fatty-acid-CoA ligase	-78	ATGACAC-(14)-ATGGTAG	4.44
RHA1_RS31935		Ammonium transporter (†)	-49	CTCAGGT-(14)-GGCTGGG	4.37
RHA1_RS31150		molybdopterin biosynthesis protein (†)	-46	CGCAGCC-(13)-CTGAGGG	4.35
RHA1_RS27660		Glutamine ABC transporter permease (†)(‡)	-213	ATCACGG-(13)-CGGCGAT	4.35
RHA1_RS37765		Acyl-CoA synthetase	-189	CTGACGC-(14)-GCGTCAG	4.31
RHA1_RS44255		Malate decarboxilase (oxalacetate decarboxilating) NADP-dependent	-36	ATCGCCC-(13)-CGAGGAG	4.30
RHA1_RS00150-RHA1_RS00140	*--atf1-atf2*	HP-DGAT-DGAT(*)	-82	ATCGCAG-(14)-CGGGGAA	4.25

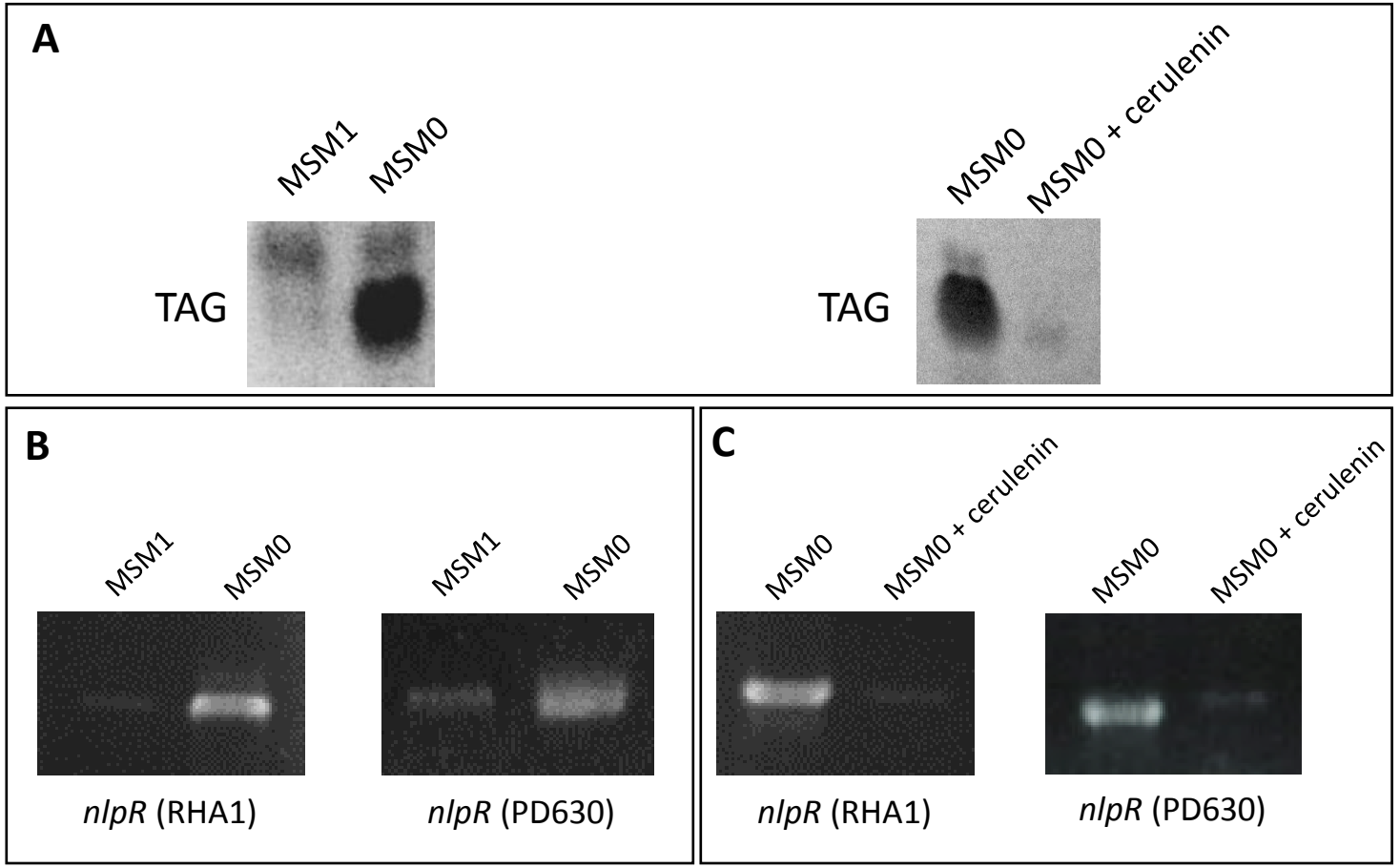
(*)Occurrence of a binding motif also with MEME/MAST tools. (†)Genes upregulated under nitrogen-limiting conditions in RHA1 strain as Amara *et al.*, 2016. (††)Genes upregulated under nitrogen-limiting conditions in PD630 strain as Chen *et al.*, 2015. Discussed genes in the text are in **bold**.

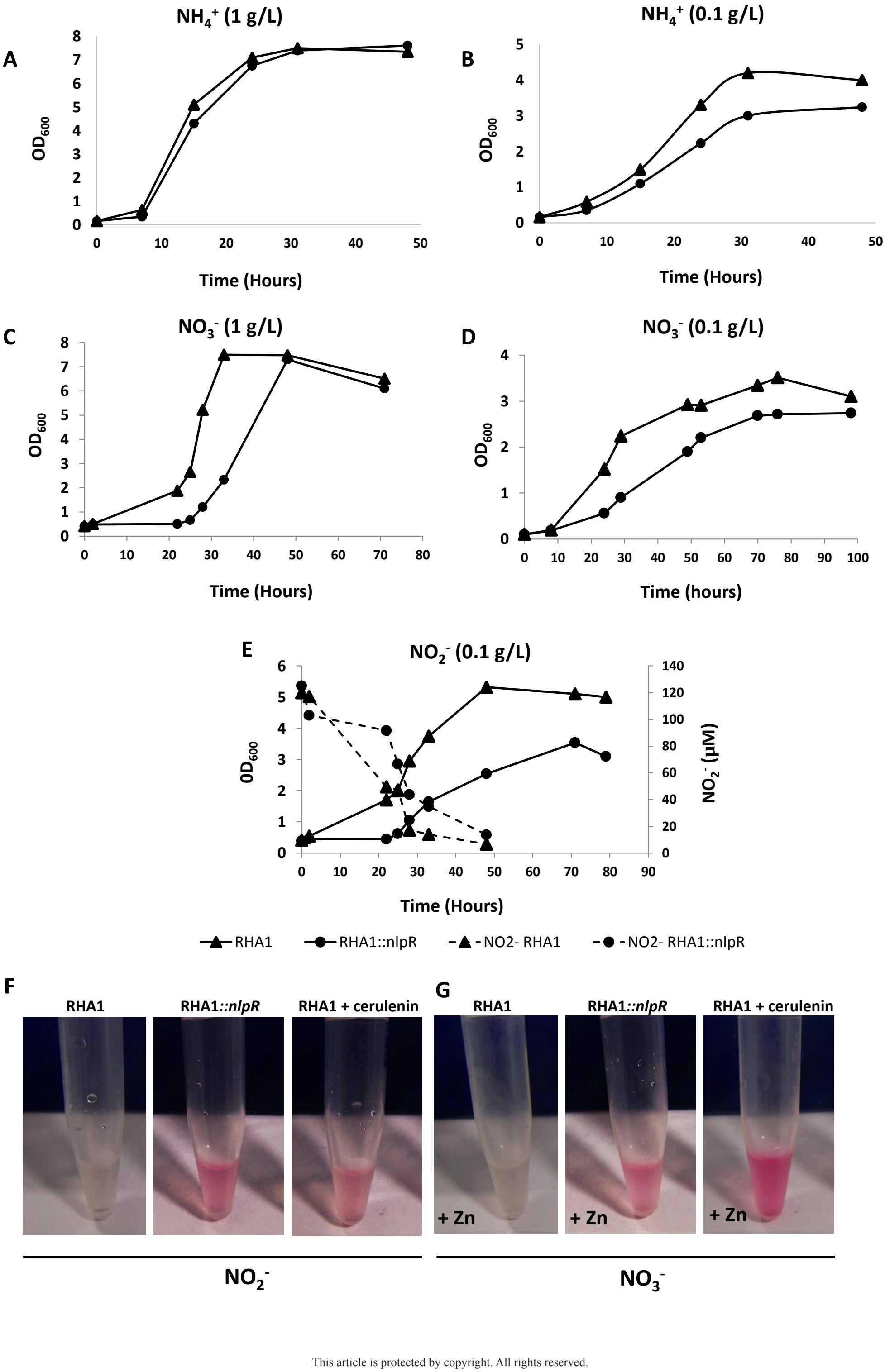
Table 3: Strains and plasmids used in this study

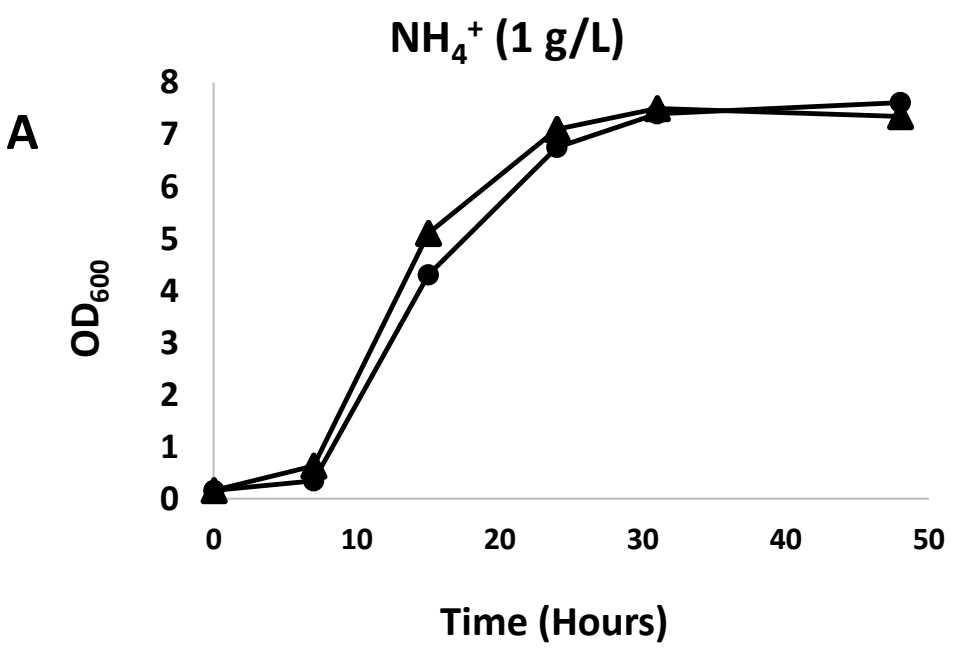
Strain/plasmid	Description	Reference /Source
Strains		
E. coli		
DH5α	*E. coli K-12 F⁻ lacU169 (ϕ80lacZΔM15) endA1 recA1 hsdR17 deoR supE44 thi-1-l2 gyrA96 relA1*	Hanahan,1983
JM109	*endA1 glnV44 thi-1 relA1 gyrA96 recA1 mcrB⁺ Δ(lac-proAB) e14- [F' traD36 proAB⁺ lacI^q lacZΔM15] hsdR17($r_K^- m_K^+$)*	Promega
BL21(DE3)	*E. coli str. B F– ompT gal dcm lon hsdSB(rB–mB–) λ(DE3 [lacI lacUV5-T7p07 ind1 sam7 nin5]) [malB+]K-12(λS)*	Novagen
BL21(DE3) pET28/*nlpR*	BL21(DE3) strain carrying the expression plasmid pET28/*nlpR*	This study
Rhodococcus		
PD630	*R. opacus* PD630, Parental strain	DSM 44193.
RHA1	*R. jostii* RHA1, Parental strain	Seto el al. 1995
602	*R. jostii* 602, Parental strain	Silva *et al.,* 2010
RHA1::*nlpR*	Mutant strain obtained by *nlpR* disruption, Km^R; derivative of RHA1	This study
RHA1::*nlpR*C	RHA1::*nlpR* strain complemented with pTipQC2/*nlpRt*; Cm^R, $Thio^R$	This study
PD630::*nlpR*	Mutant strain obtained by *nlpR* disruption, Km^R; derivative of PD630	This study
602::*nlpR*	Mutant strain obtained by *nlpR* disruption, Km^R; derivative of 602	This study
RHA1 pTipQC2	RHA1 derivative carrying pTip-QC2 plasmid, used as control strain; Cm^R	This study
RHA1 pTipQC2/*nlpR*	RHA1 derivative carrying pTip-QC2 /*ro06368t*; Cm^R; $Thio^R$	This study
PD630 pTipQC2	PD630 derivative carrying pTip-QC2 plasmid, used as control strain; Cm^R	This study
PD630 pTipQC2/*nlpR*	PD630 derivative carrying pTip-QC2 /*ro06368t*; Cm^R; $Thio^R$	This study
Plasmids		
pGEM-T-easy vector	Linear plasmid used for cloning purposes; Ap^R	Promega
pGEM-T-easy/"*nlpR*"	pGEM-T-easy derivative carrying a partial 793 bp region of *ro06368*; Ap^R	This study
pGEM-T-easy/"*nlpR*"-ΩKm	pGEM-T-easy/"*nlpR*" carrying a kanamycin cassette in *Pst*I site; Ap^R, Km^R	This study
pGEM-T-easy/*nlpR*	pGEM-T-easy derivative carrying the complete 1158 bp of *ro06368*; Ap^R	This study
pTip-QC2	Expression vector for *Rhodococcus* with P*tipA* promoter, *repAB* (pRE2895); Cm^R	Nakashima *et al.,* 2004
pTip-QC2/*nlpR*	pTip-QC2 carrying the complete *ro06368* gene under control of P*tipA*; Cm^R	This study
pET28	Vector for expression of genes under control of the T7 promoter; Ap^R	Novagen
pET28/*nlpR*	pET28 carrying the complete *ro06368* gene; Ap^R	This study

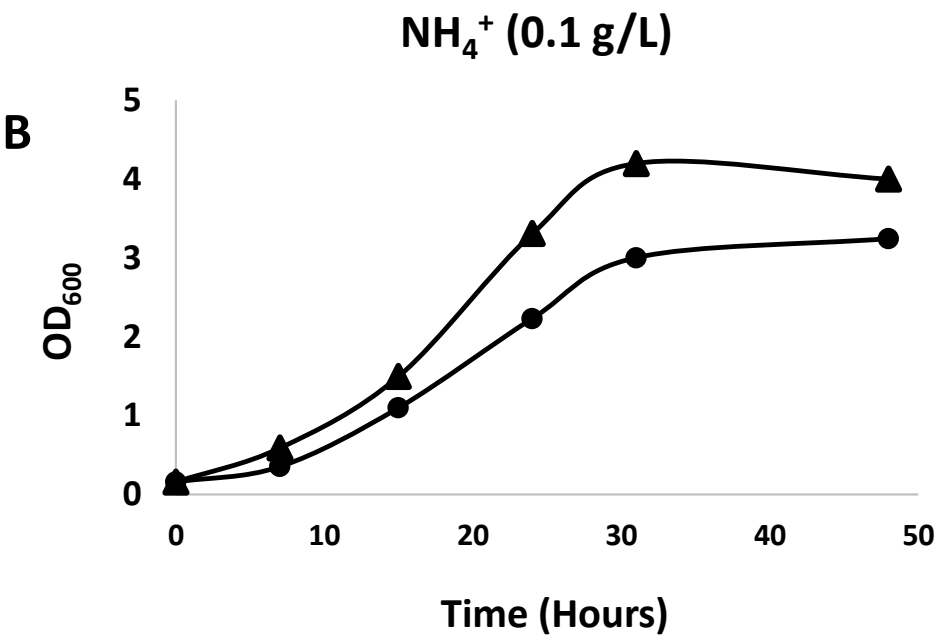


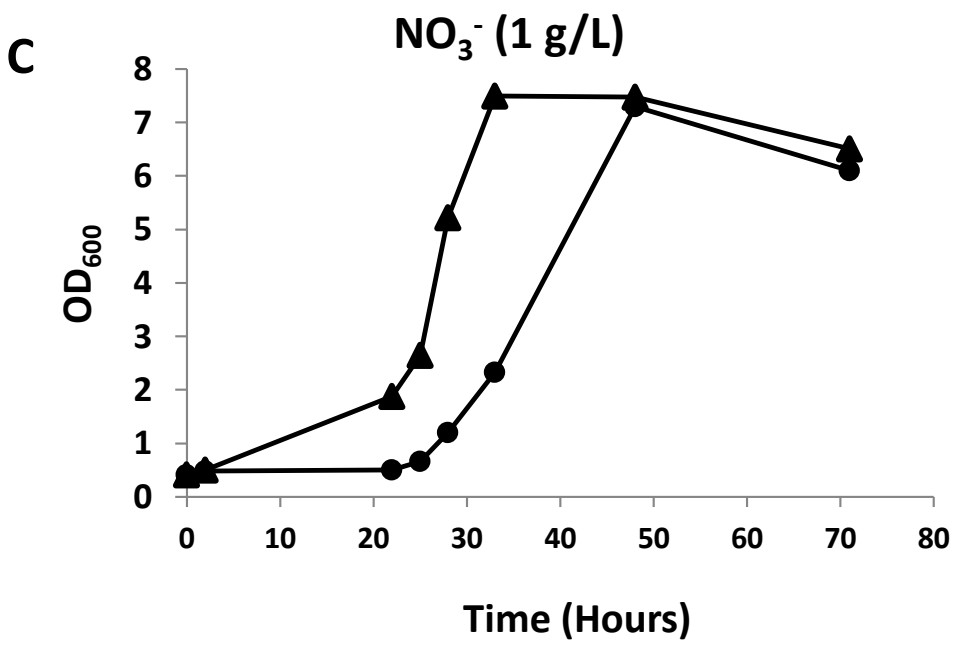


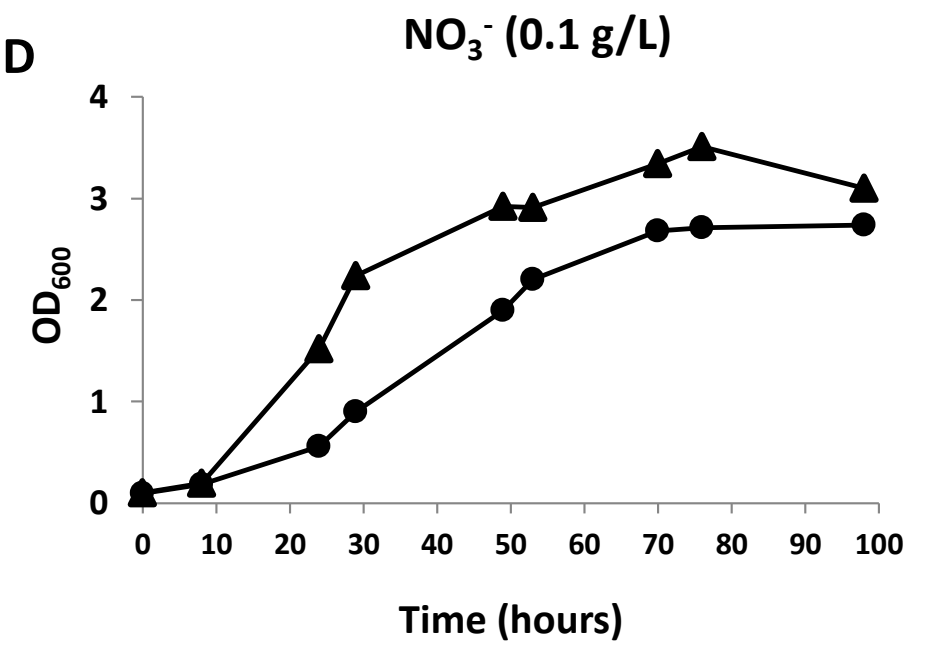


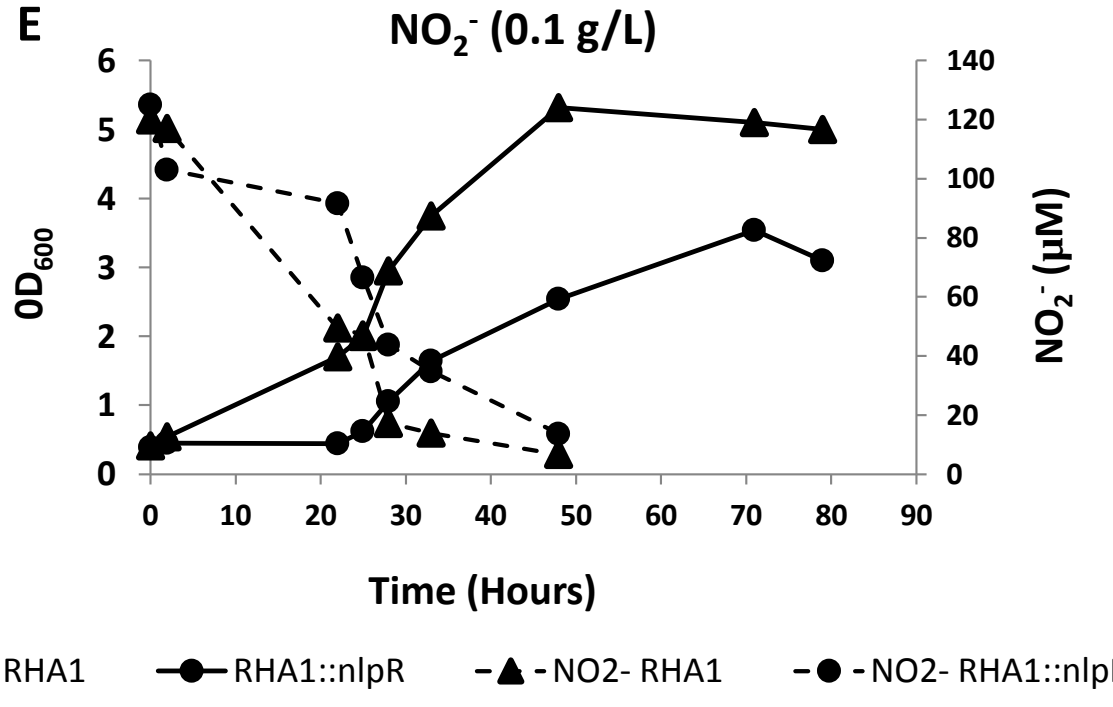


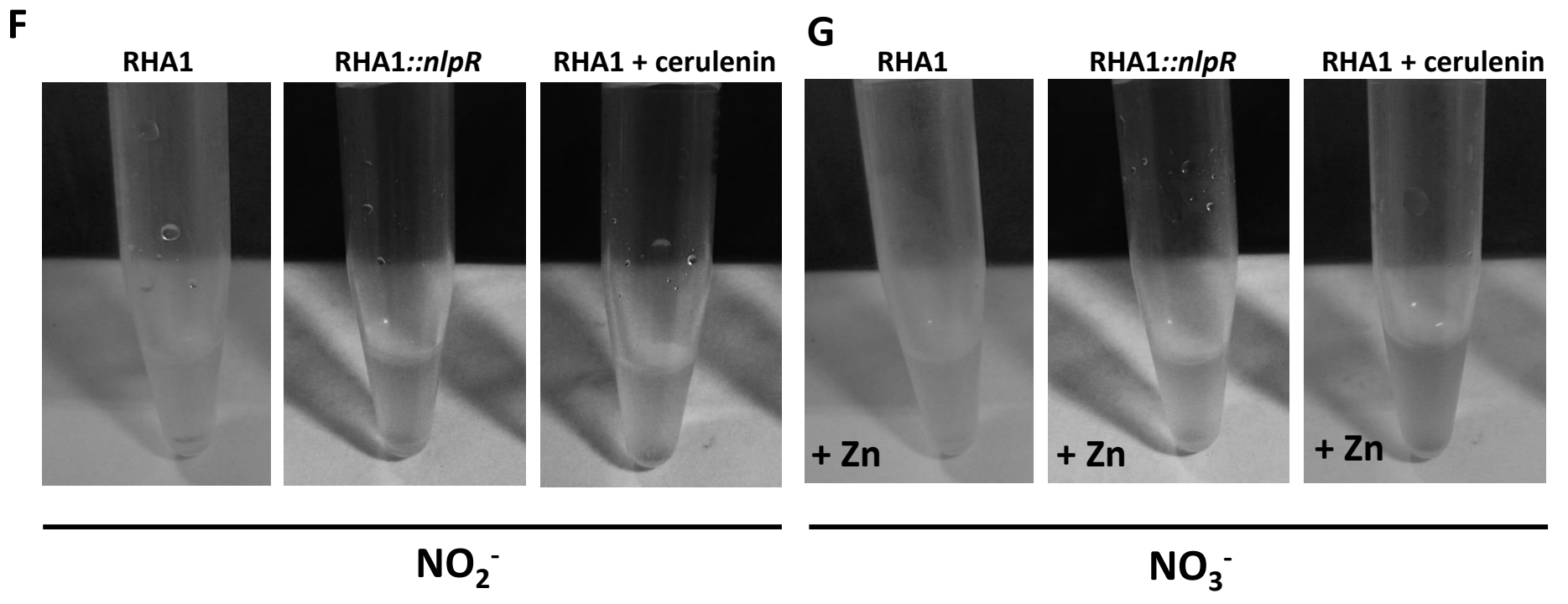


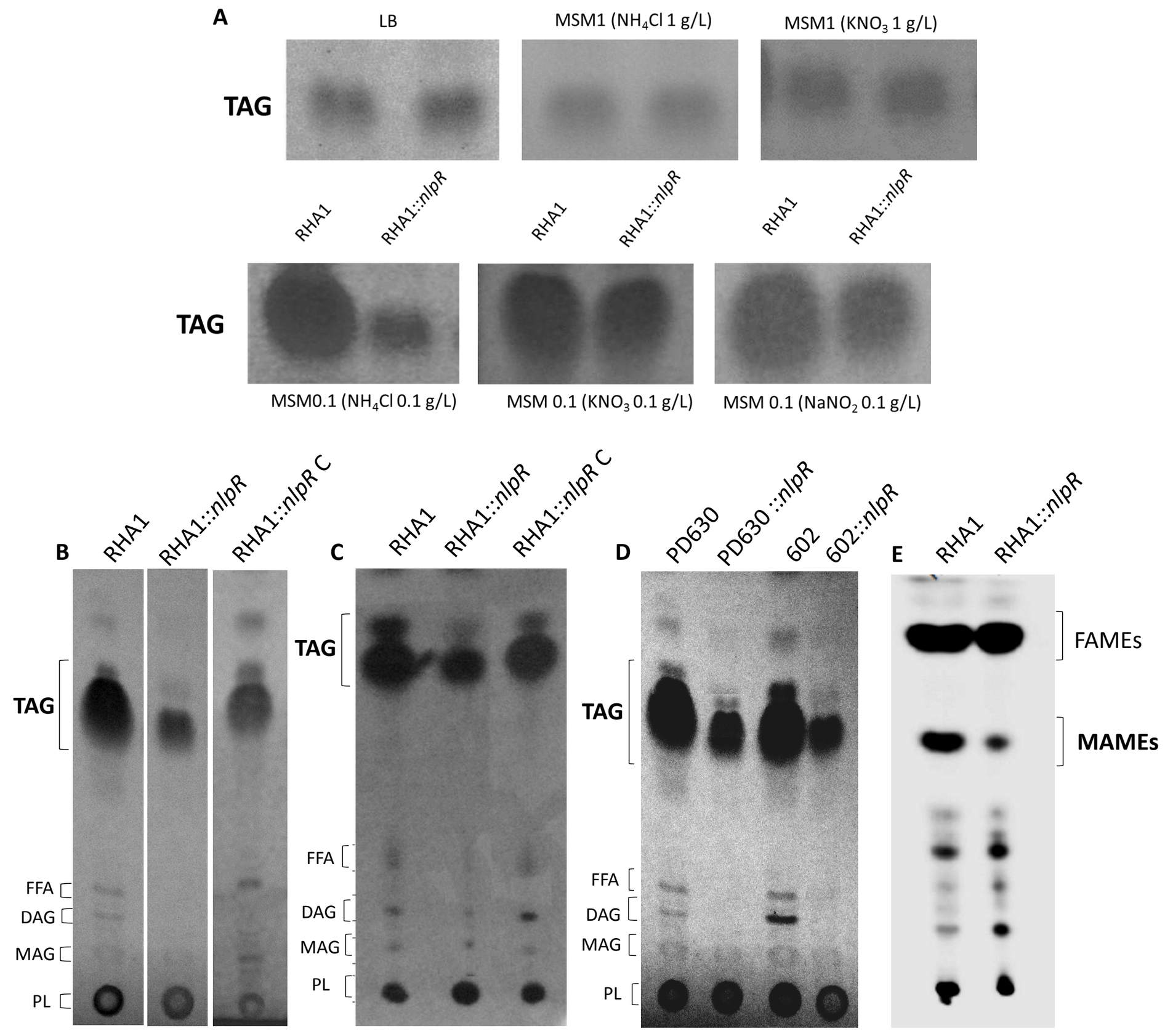


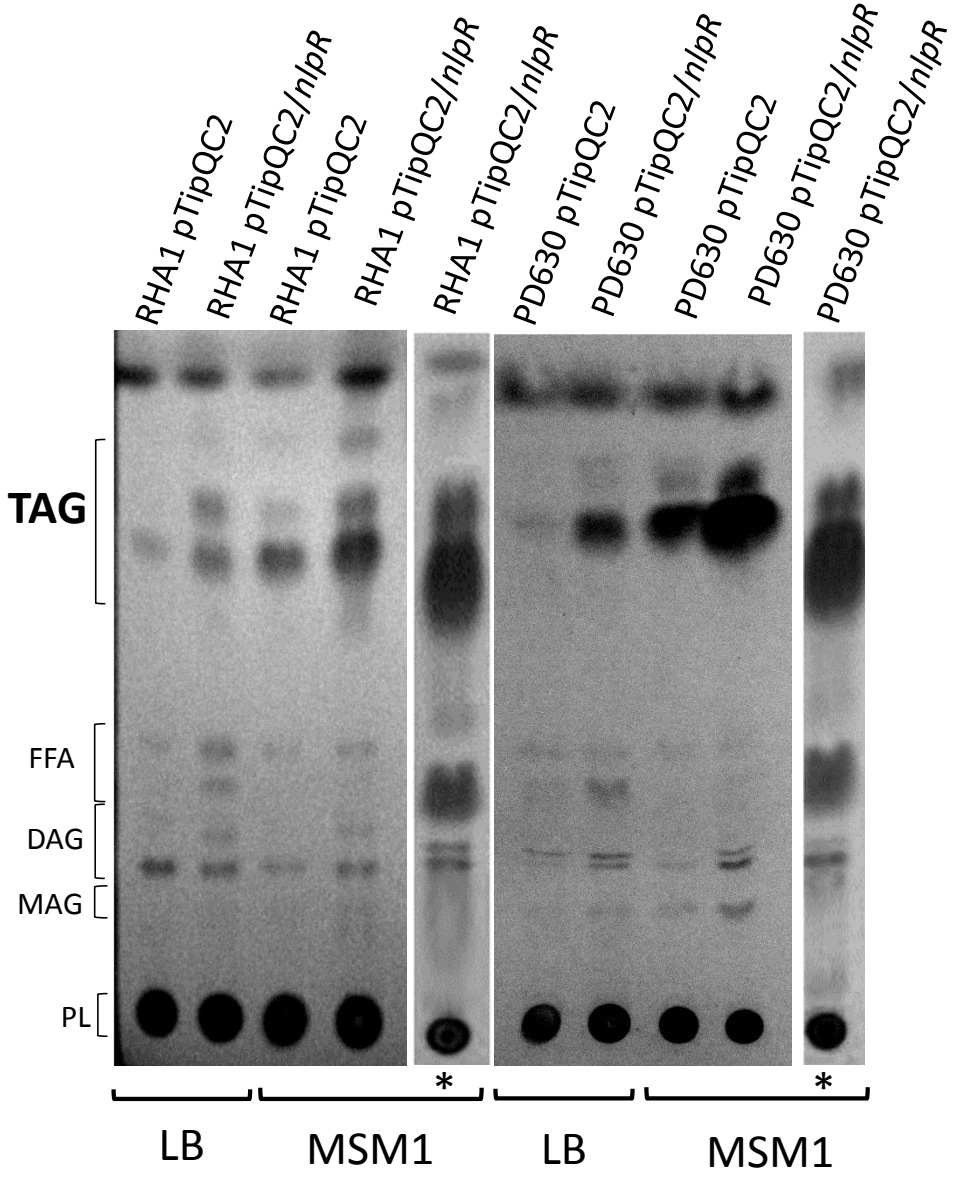








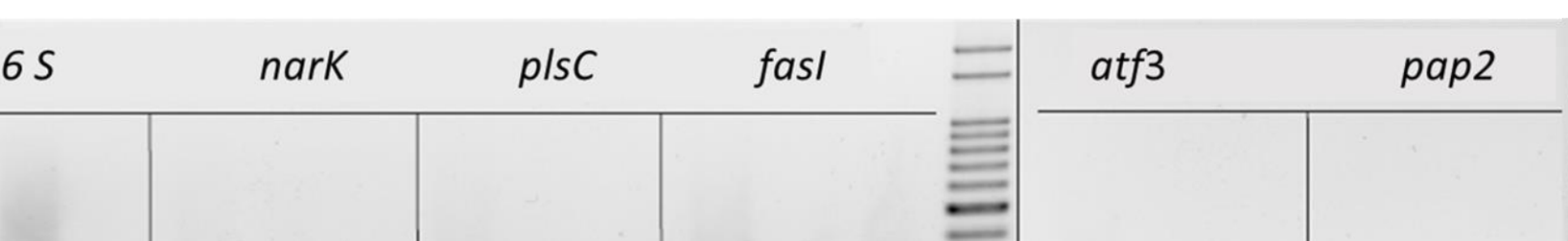




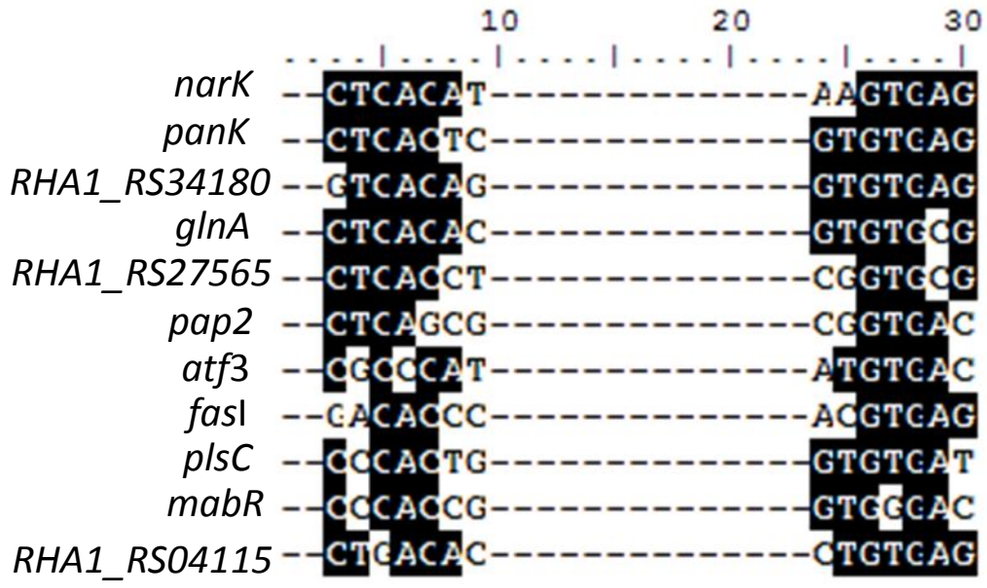
A

```
SCO_NARK   CTCTCACCCCGCGCCCGTGCCGGGTGAGAGCGAC        SCO_NARK   CTCACCCNNNNNNNNNNNNNNGTGAGAG
SCO_NIRB   CCCTCACACGCGCCGGCCGGCCCCGCTGAGGGCA        SCO_NIRB   CTCACACNNNNNNNNNNNNNNCGCTGAG
SCO_NASA   CCCTCACCGCGGCCCGGCGGGGGCGGGGAGGTCG        SCO_NASA   CTCACCGNNNNNNNNNNNNNNCGGGGAG
MTU_NARK   CTCTCACACAGCCGGTCGCCAG-GTGTGAGCACA        MTU_NARK   CTCACACNNNNNNNNNNNNNN-GTGTGAG
MTU_NIRB   CCGTCACACGCTGGCGGTAACC-GTGTGATCGAA        MTU_NIRB   GTCACACNNNNNNNNNNNNNN-GTGTGAT
RHO_NARK   ACCTCACATCGGCCCAGCTGAA-AAGTGAGGATC        RHO_NARK   CTCACATNNNNNNNNNNNNNN-AAGTGAG
             *******               *******                       *******               *******
```

B

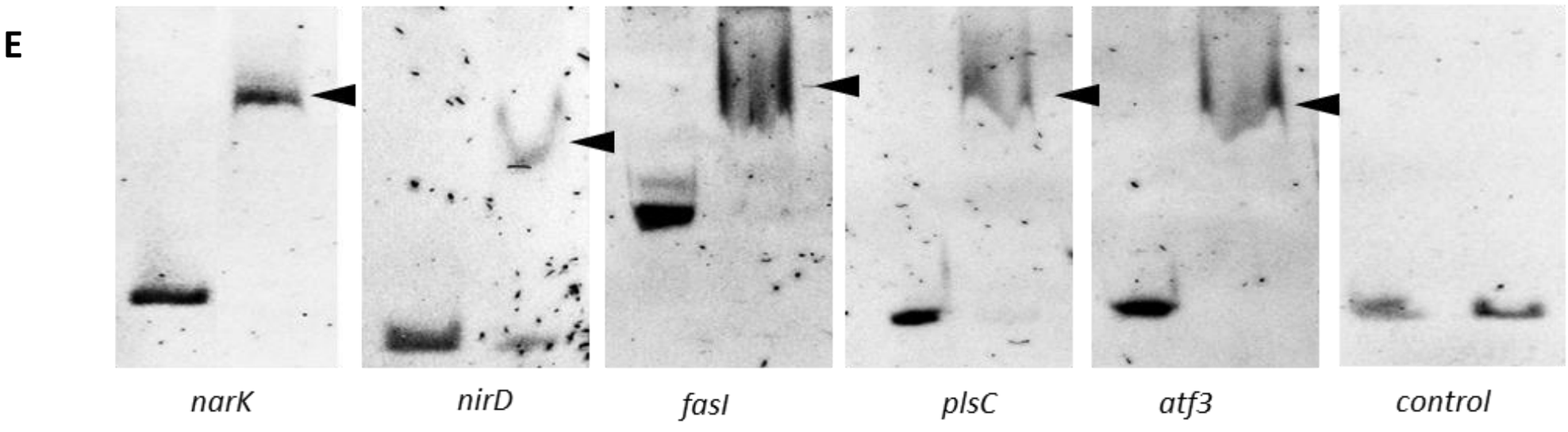


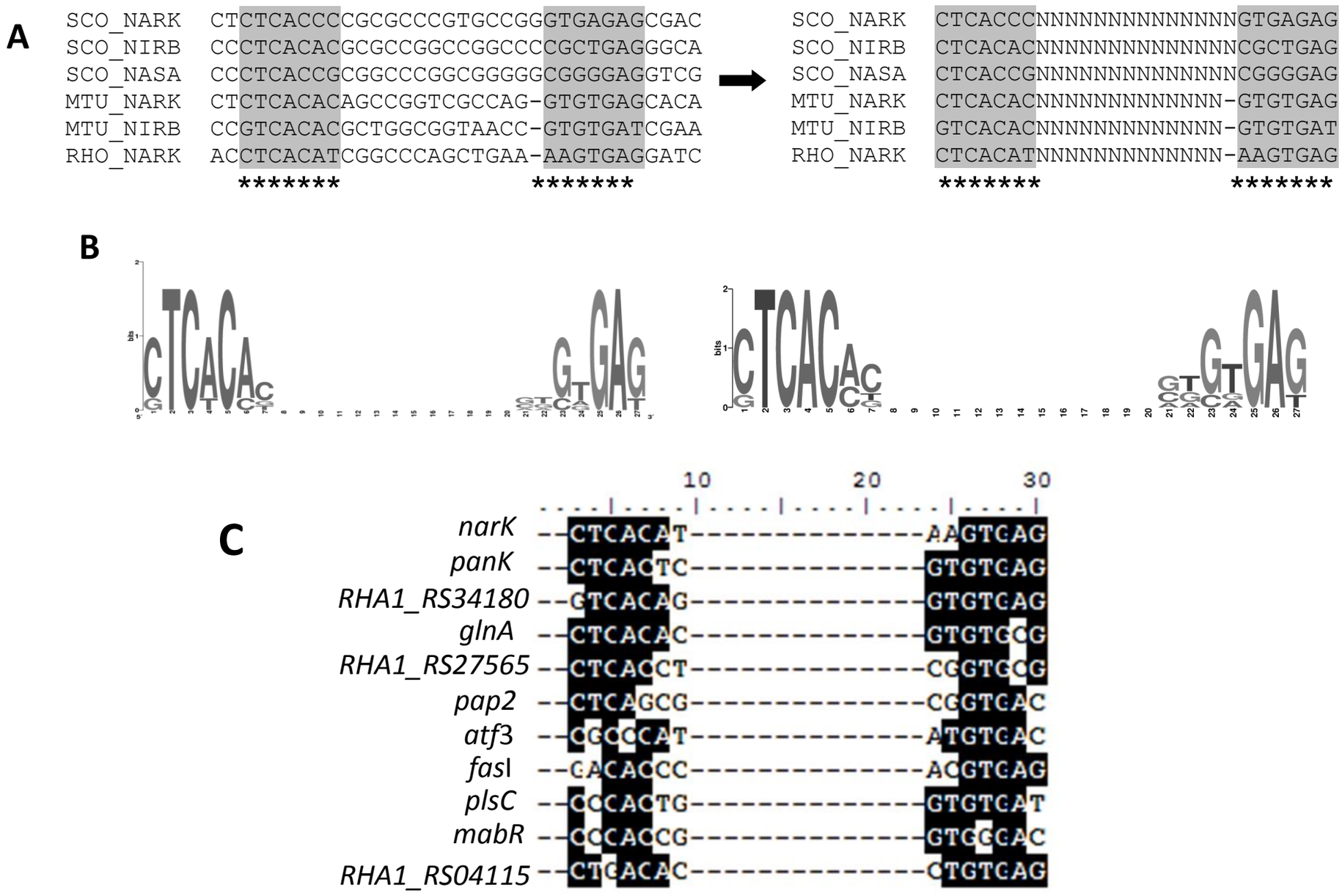
C

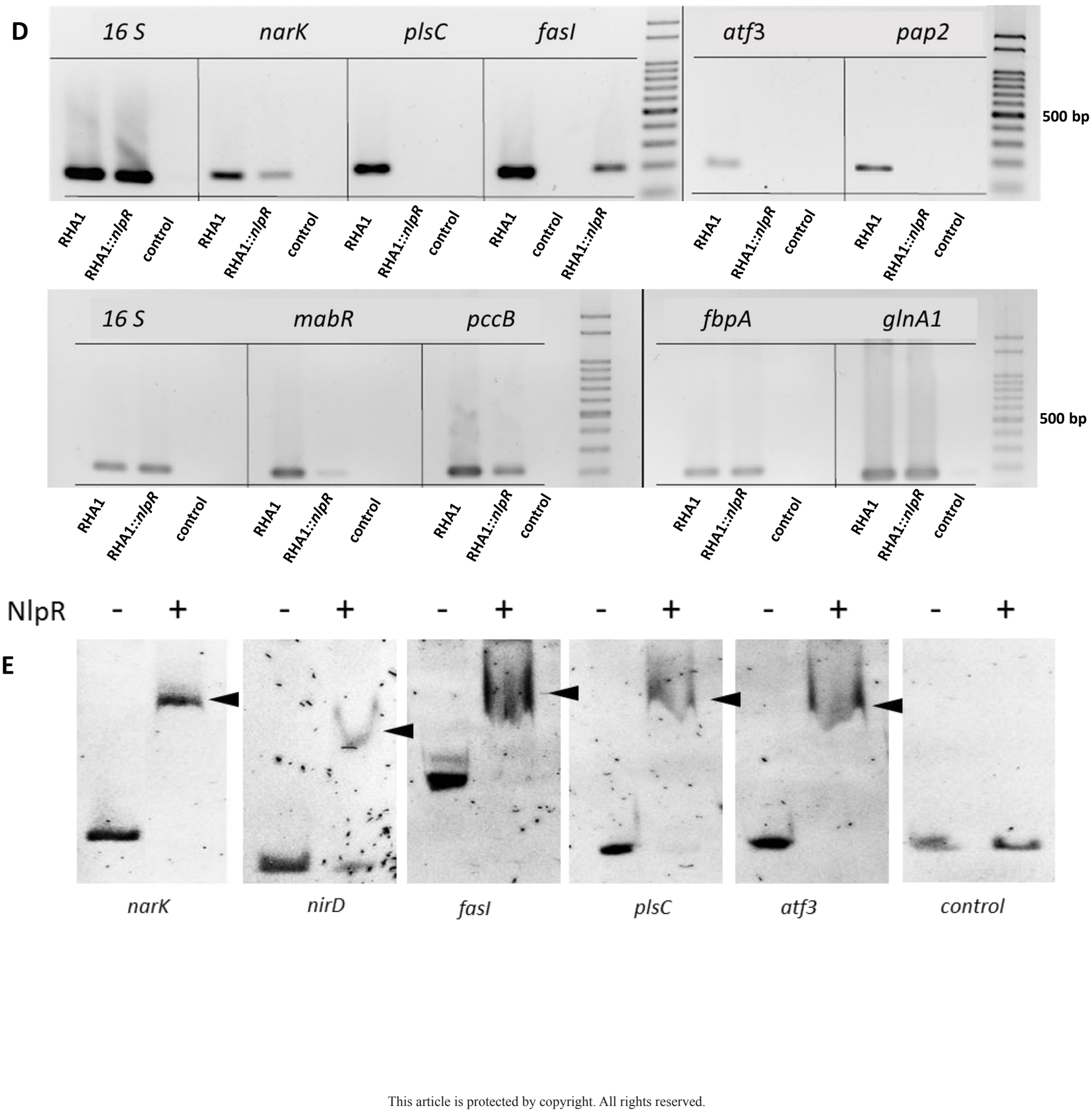


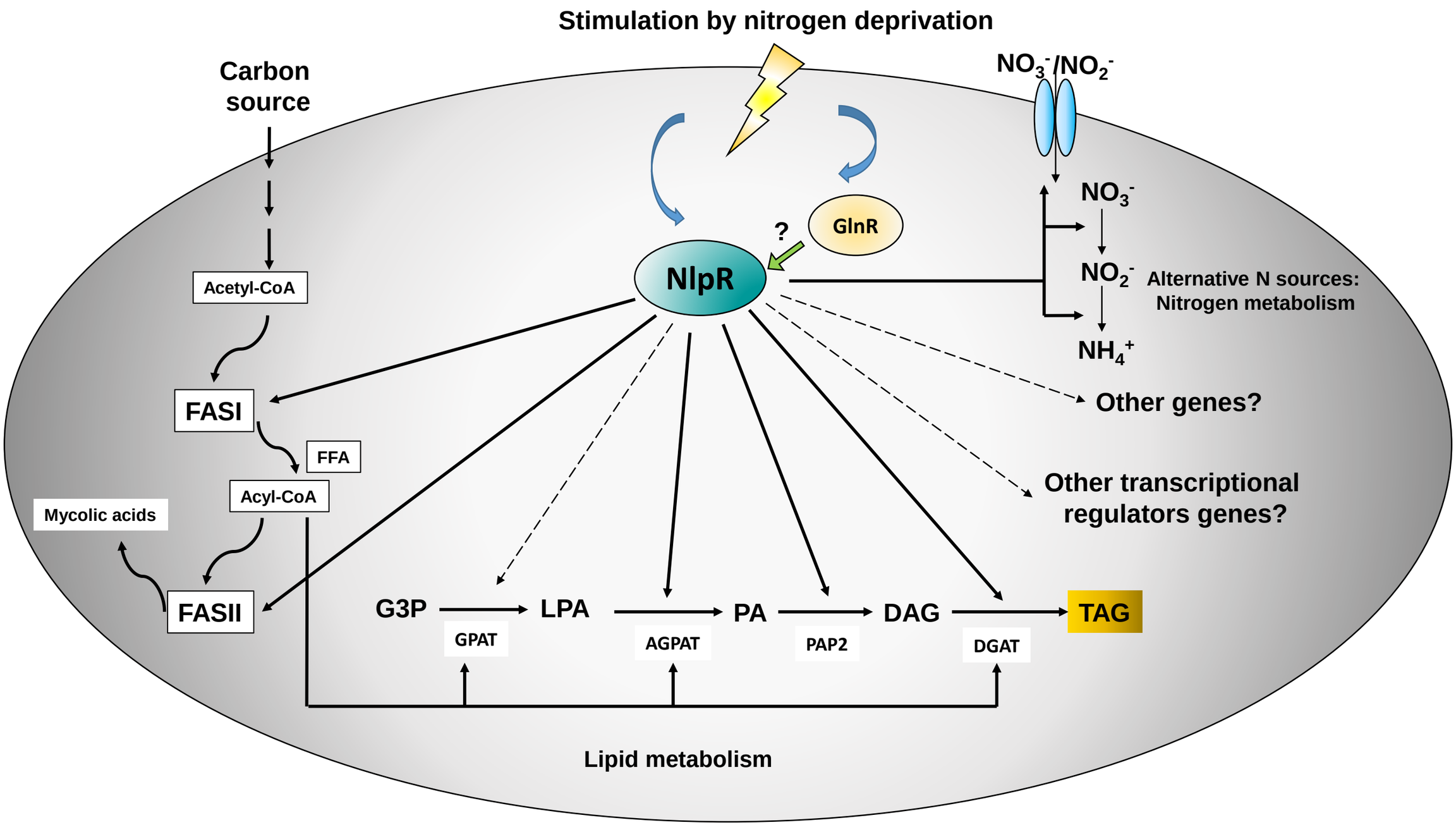
D

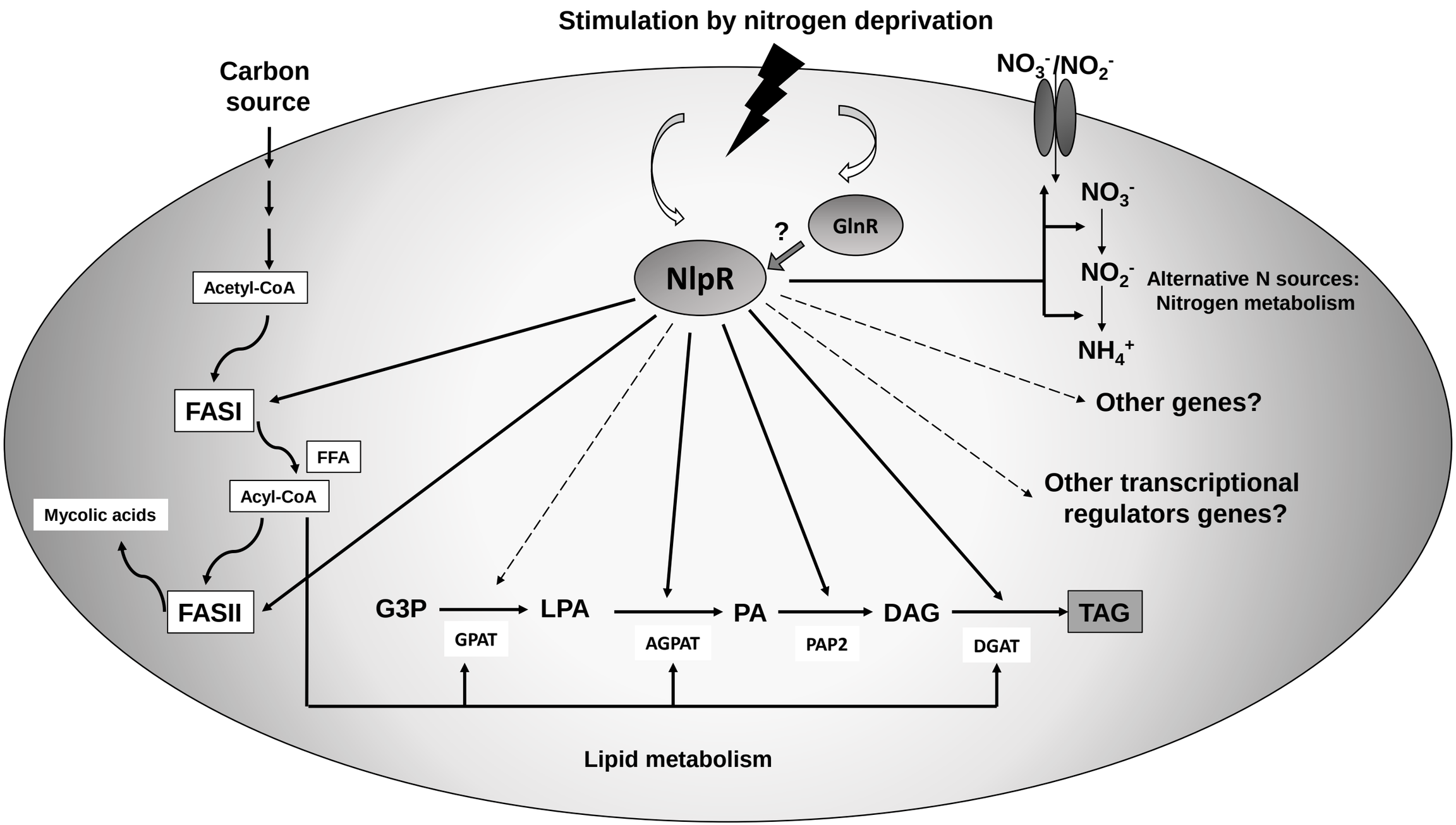
E











Abbreviated Summary

The ability to accumulate significant amounts of triacylglycerols by rhodococci requires a special metabolic network involving concerted reactions and pathways, and a specific regulatory circuit for the control and coordination of the process. In this study, we identified a transcriptional regulator that contributes to control of nitrogen metabolism and the modulation of lipid biosynthesis and triacylglycerol accumulation in *R. jostii* RHA1 under nitrogen starvation conditions.